

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2003, 17:45:31 ; Search time 10243 Seconds
(without alignments)
11562.371 Million cell updates/sec

Title: US-09-981-151A-7

Perfect score: 2895

Sequence: 1 cgcctccggatgaagcccg.....cccggggcttctccctgccca 2895

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2898711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_ats.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_cm.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_ats.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2895	100.0	2895	6	AX574596 Sequence
2	2675.4	92.4	2997	6	AX574590 Sequence
3	1877.8	64.9	3227	9	AX315734 Homo sapi
4	1877.8	64.9	3675	6	AX319852 Sequence
5	1877.8	64.9	5610	6	BD170080 Novel ADA
6	1877.8	64.9	8435	6	BD170083 Novel ADA
7	1877.8	64.9	8503	6	BD170084 Novel ADA
8	1874.6	64.8	3675	6	AR229456 Sequence
9	1874.6	64.8	3675	6	BD170564 Novel pro
10	1874.6	64.8	4042	6	AR229457 Sequence
11	1795.6	62.0	2433	6	AX574592 Sequence
12	1695	58.5	2902	6	AX574594 Sequence
13	1634.8	56.5	4234	6	BD170079 Novel ADA
14	1634.8	56.5	4234	9	AB095949 Homo sapi
15	1578.4	54.5	7668	6	BD170085 Novel ADA
16	1381.8	47.7	1668	6	AX481380 Sequence
17	1305.4	45.1	3402	10	BC034739 Mus muscu
18	708	24.5	3666	6	BD170498 Novel agg
19	706.4	24.4	2805	6	AX319854 Sequence
20	706.4	24.4	3445	9	AX311903 Homo sapi
21	663	22.9	1476	6	AR229455 Sequence
22	417.6	14.4	545	6	BD170082 Novel ADA
23	320.2	11.1	422	6	BD170081 Novel ADA
24	311.6	10.8	104776	9	AC022424 Homo sapi
25	310	10.7	67468	2	AC139148 Homo sapi
26	296	10.2	187084	9	AC010269 Homo sapi
27	259.2	9.0	82998	2	AC022970 Homo sapi
28	227	7.8	2879	6	BD079104 Novel met
29	227	7.8	2879	6	BD079092 Novel met
30	215.4	7.4	5338	6	BD079104 Novel met
31	215.4	7.4	5338	6	BD079094 Novel met
32	215.4	7.4	5338	6	BD079106 Novel met
33	211.8	7.3	5353	6	AX128433 Sequence
34	210.6	7.3	5061	6	E55265 Novel metal
35	208	7.2	3132	6	AX327754 Sequence
36	208	7.2	3312	6	AX319851 Sequence
37	208	7.2	3312	6	E55273 Novel metal
38	208	7.2	3329	6	AX327752 Sequence
39	208	7.2	3400	9	AF163762 Homo sapi
40	206.4	7.1	3377	6	AX467546 Sequence
41	206.2	7.1	3207	6	AX327747 Sequence
42	206.2	7.1	3402	6	AX338539 Sequence
43	206.2	7.1	3403	6	AX327745 Sequence
44	206.2	7.1	3403	6	AX338537 Sequence
45	204.4	7.1	2445	6	AX327757 Sequence

ALIGNMENTS

RESULT 1
AX574596
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX574596
Sequence 7 from Patent WO0233087.
AX574596
AX574596.1 GI:27551852

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae, Homo.

Edinger, S., Gerlach, V., Macdougall, J.R., Malyankar, U.M.,
Smithson, G., Millet, I., Peyman, J.A., Stone, D.J., Gunther, E.,
Ellerman, K., Shinkets, R.A., Padigaru, M., Guo, X., Patturajan, M.,

linear PAT 07-JAN-2003

Taupier R.J., Burgess, C.E., Zerhusen, B.D., Kekuda, R., Spytek, K.A.,
Gangoli, E.A., Fernandes, E.R. and Gorman, L.
Proteins and nucleic acids encoding same
Patent: WO 0233087-A 7 25-APR-2002;
Curagen Corporation (US)

FEATURES
Location/Qualifiers

1..2895
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 699 a 760 c 787 g 549 t

Query Match 100.0%; Score 2895; DB 6; Length 2895;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2895; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CGCTCTGGATGAAGCCCGCGCGGATGCGGGCTTGGCGCGTGTGGATGCTG 60
DB 1 CGCTCTGGATGAAGCCCGCGCGGATGCGGGCTTGGCGCGTGTGGATGCTG 60
QY 61 TTGGCGAGGTGGCGGAGCGGTGAGTCCCGGCGCTCCACACGCGGAAACCGCGG 120
DB 61 TTGGCGAGGTGGCGGAGCGGTGAGTCCCGGCGCTCCACACGCGGAAACCGCGG 120
QY 121 TCGGACAGCTGGAGGCGAGTCCCGCGCGCTCTCTCCGCGGACCCCGCTCCAC 180
DB 121 TCGGACAGCTGGAGGCGAGTCCCGCGCGCTCTCTCCGCGGACCCCGCTCCAC 180
QY 181 GCGATGTCGCGCTGTTTCCGAGGACCTGCGTGCGCATGGACCCCGAGCGAGC 240
DB 181 GCGATGTCGCGCTGTTTCCGAGGACCTGCGTGCGCATGGACCCCGAGCGAGC 240
QY 241 GCCTGGAGCCCGAGCTCCCGCTCTCTCCACCGCGAGCGCGCGCTGATGGA 300
DB 241 GCCTGGAGCCCGAGCTCCCGCTCTCTCCACCGCGAGCGCGCGCTGATGGA 300
QY 301 AAAGCGGGGACATGATGATGAAACCATGTTCTCAGCAAACTAACACAGGAACA 360
DB 301 AAAGCGGGGACATGATGATGAAACCATGTTCTCAGCAAACTAACACAGGAACA 360
QY 361 GAAACCAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 GAAACCAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 CACAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 CACAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GCGTGTCCGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GCGTGTCCGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 ATGGATCTGAGGACTTCCAGCAGCTAGTGGCTCTGGCTTTATTTGTCAGAGCTTGGGA 600
DB 541 ATGGATCTGAGGACTTCCAGCAGCTAGTGGCTCTGGCTTTATTTGTCAGAGCTTGGGA 600
QY 601 AAGACGACCTAGTCTGTCAGACTTACCGCAGAGGACTCTGTTCTATATCAGAGC 660
DB 601 AAGACGACCTAGTCTGTCAGACTTACCGCAGAGGACTCTGTTCTATATCAGAGC 660
QY 661 TCTTTGCGATCACACAGAACTCCCATCGATGAGGGAAGTCTGTGAGGCTCCACT 720
DB 661 TCTTTGCGATCACACAGAACTCCCATCGATGAGGGAAGTCTGTGAGGCTCCACT 720
QY 721 GCGACTCTGAGGCTCTGCAAGCTCAGAAATGTCCTCCGAGACAGTGTGACTTCCGTGT 780
DB 721 GCGACTCTGAGGCTCTGCAAGCTCAGAAATGTCCTCCGAGACAGTGTGACTTCCGTGT 780
QY 781 GCTCAGTGTCCGAGCACAACAGCAGAGGATTCAGAGCGGCGCACTACAGTGGAGCCT 840
DB 781 GCTCAGTGTCCGAGCACAACAGCAGAGGATTCAGAGCGGCGCACTACAGTGGAGCCT 840

```

```

QY 841 TACACTCAAGTAGAAGCCGACTTATGCAAACTCTACGTATCCGAGAAGATTGATTTC 900
DB 841 TACACTCAAGTAGAAGCCGACTTATGCAAACTCTACGTATCCGAGAAGATTGATTTC 900
QY 901 TTCTTTTCTTTGTCAAATAAAGTCAAGATGGGACTCCATGCTCGGAGGATACCGTAAT 960
DB 901 TTCTTTTCTTTGTCAAATAAAGTCAAGATGGGACTCCATGCTCGGAGGATACCGTAAT 960
QY 961 GTTTGTATAGATGGATATCTGAGCTCAGTGTGGTGTCCACATCTCGGCACATGCCCCAG 1020
DB 961 GTTTGTATAGATGGATATCTGAGCTCAGTGTGGTGTCCACATCTCGGCACATGCCCCAG 1020
QY 1021 CCTCCGAGGAGACCTCTTCATCTTCCAGATGAGTATAGTCTTGTCTACGGCATAG 1080
DB 1021 CCTCCGAGGAGACCTCTTCATCTTCCAGATGAGTATAGTCTTGTCTACGGCATAG 1080
QY 1081 CGCTCTCTTCTGAGTCCCATAGAAATGAAGAACTGAACTGGGAGACCTTGGTGTGGTTC 1140
DB 1081 CGCTCTCTTCTGAGTCCCATAGAAATGAAGAACTGAACTGGGAGACCTTGGTGTGGTTC 1140
QY 1141 GACAAAAGATGATGCAAAACCATGGCCATGAAATAATACCACTTACGTCTCAGATA 1200
DB 1141 GACAAAAGATGATGCAAAACCATGGCCATGAAATAATACCACTTACGTCTCAGATA 1200
QY 1201 CTCAACATGGTATCTGCTTTTATTCAAGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 CTCAACATGGTATCTGCTTTTATTCAAGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 GACCAACCATCTTACTGACTGGCTTCAACATTTGCCATGAGTCTGGACAACTTTGGC 1320
DB 1261 GACCAACCATCTTACTGACTGGCTTCAACATTTGCCATGAGTCTGGACAACTTTGGC 1320
QY 1321 ACTTTGGGATTTGACCCATAGTGAATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 ACTTTGGGATTTGACCCATAGTGAATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 GAAGATCAGGCTCTTGGACTGGCTTCAACATTTGCCATGAGTCTGGACAACTTTGGC 1440
DB 1381 GAAGATCAGGCTCTTGGACTGGCTTCAACATTTGCCATGAGTCTGGACAACTTTGGC 1440
QY 1441 ATGATTCATGATGAGGAGGAGGACATGTTAAAGTCCGAGGACATCATGTCCCT 1500
DB 1441 ATGATTCATGATGAGGAGGAGGACATGTTAAAGTCCGAGGACATCATGTCCCT 1500
QY 1501 ACATTTGAGGACCAATGAGTCTTCTCTGCTGCTACCTTCCAGCCCGAGTATCTACAC 1560
DB 1501 ACATTTGAGGACCAATGAGTCTTCTCTGCTGCTACCTTCCAGCCCGAGTATCTACAC 1560
QY 1561 AAATTTCTAGCAGCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCTTGTGAAGGA 1620
DB 1561 AAATTTCTAGCAGCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCTTGTGAAGGA 1620
QY 1621 TACAAGTATCTGAGAAATTTGCCAGGAGATTTATGATGCAAAACACACAGTGCAGTGG 1680
DB 1621 TACAAGTATCTGAGAAATTTGCCAGGAGATTTATGATGCAAAACACACAGTGCAGTGG 1680
QY 1681 CAGTTTCGAGAGAAAGCCAGCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1740
DB 1681 CAGTTTCGAGAGAAAGCCAGCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1740
QY 1741 CTGTGTGCTCCATCTGATTTGGAAGGAAATGTGAGACTTAAATTTATGCCAGCAGAGAAGGC 1800
DB 1741 CTGTGTGCTCCATCTGATTTGGAAGGAAATGTGAGACTTAAATTTATGCCAGCAGAGAAGGC 1800
QY 1801 ACAATTTGTGGCATGACATGTGTGCGGGAGGACAGTGTGTGAATATGTGTATGAA 1860
DB 1801 ACAATTTGTGGCATGACATGTGTGCGGGAGGACAGTGTGTGAATATGTGTATGAA 1860
QY 1861 GGCCCCAAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
DB 1861 GGCCCCAAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
QY 1921 ACCTCGGAGGGGGAGTATCTCATAGAGTCTGCTCTGCAACCAACCCCAAGCCATCGCAT 1980

```


QY	2755	TCGTAGTCTCGTCTCGGAGGGGTAGTGCCTTCCAGTCTGCTCTCGTCTCGAGGCA	2814
Db	2857	TCGTAGTCTCGTCTCGGAGGGGTAGTGCCTTCCAGTCTGCTCTCGTCTCGAGGCA	2916
QY	2815	GCATGTGACGCTTCAGCCACTCGGTACATTCGACCTGGCCCTTCTTGAATCCCTAATGAGCA	2874
Db	2917	GCATGTGACGCTTCAGCCACTCGGTACATTCGACCTGGCCCTTCTTGAATCCCTAATGAGCA	2976
QY	2875	GCCCGGGGCTTCTCCCTGCCA	2895
Db	2977	GCCCGGGGCTTCTCCCTGCCA	2997
RESULT 3			
LOCUS	HS315734	3227 bp	mRNA linear PRI 01-MAR-2002
DEFINITION	Homo sapiens mRNA for metalloprotease disintegrin 16 (ADAMTS16 gene).		
ACCESSION	AJ315734		
VERSION	AJ315734.1	GI:19171177	
KEYWORDS	ADAMTS16 gene; disintegrin; metalloprotease; thrombospondin.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Cal, S., Obaya, A.J., Llamazares, M., Garabaya, C., Quesada, V. and Lopez-Otin, C.		
TITLE	Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains		
JOURNAL	Gene 283 (1-2), 49-62 (2002)		
MEDLINE	21856482		
PUBMED	11867212		
REFERENCE	2 (bases 1 to 3227)		
AUTHORS	Cal, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-JUN-2001) Cal S., Biochemistry and Molecular Biology, University of Oviedo, Campus del Cristo, Asturias. 33006, SPAIN		
FEATURES	Location/Qualifiers		
source	1..3227	/organism="Homo sapiens"	
gene	/mol_type="mRNA"	/db_xref="taxon:9606"	
CDS	1..3227	/gene="ADAMTS16"	
	1..3219	/gene="ADAMTS16"	
	/codon_start=1	/product="metalloprotease disintegrin 16 with thrombospondin type I motif"	
	/protein_id="CAC86015.1"	/db_xref="GI:19171178"	
	/translation="MKPRARWGLAALWMLAQVAEQAPACAMGPAAPGSPVPR PPPAPRPMHEGEYDLVSAIEVDHGDYVSHIIMHQRRAVAVSEVSHLRUK GPRDHPMDITLSSVAPGFIQVTLGKTGKSVQVLPPEDFCFYQSGLSRHSRNSVA LSTCGQLSGMIRTEADYFLRPLSHLWKLGRAGQSSPSHVLYKRPSTEPHAPGASE VLATRWELAHOPLHSSDLRLGPOKHFCGRKKYMPQPPKEDLFLPDEYKSLR HKRSLRSHNEELNVELTVVDKMKQNHENITVYVLTILNMYSLFKDGTIGSN INTAIVGLLEDEPGLVISHADHTLSFCQWQSLGKMGKQTRHDHALLTGLDIC SWNPECDITGFAPISGMKSKYRSTINEDTGLGLAFTIAHESGHNFMIHDGEGNVC KKGSEIMSTLAFGRNWFVSPSCRSQYLHKFLSTAQAIACLADQPKPKVKEYPEKLP GELYDANTQKMGFGEKAKLMDFKKDICKALWCHRIGRKCTKMPAAEYICQHD MWCGRGCVKVGDEGPKPTGHWSWDSWSPSRTCGGVSHRSRLCTNPKPSHGKPF CEGSTPLKLNQSKQPRDSVDFRAQCAEHSNRRPRGRHVKKPKVTOVEDQDLCKLY CIAEGPFFPSLKNKYDGTFCSEDSERNVICIGICERVGCDNLGSDAEDVCGVNG NNSACTIHRGLYTHHNTQIYHMVTPSPGARSIRIYEMNVSTISLSVRNALRYILN GHWTDFWPKYKFSGTFDYRSYNEPENIATGPTNETLIVELLFQVNNPFGVAVYS PVLGLVCKVAPSPVSWGNSACSRCTCGGQASRPVQCTRRVHVDSEVPASLCPQ PAPSRQACHNSQSPKAWNSAGPWAECSTCGKWRKFAVACKSTNPSRAQLLPDAVC TSBPKPRHREACILORCHPKPKLQWLIVMSQVGLVSRERG"		
BASE COUNT	776 a	871 c	898 g 682 t

QY	1219	TTATTCAAGA-----	1229
Db	973	TTATTCAAGAAGTGAACAATAGAGGAAACATCAACATTTGCAATTGTAGTCTGATTCTT	1032
QY	1230	-----	1229
Db	1033	CTAGAAGATGAACAGCAGCAGACTGGTGATAAGTCACACGACGACACACCTTAAGTAGC	1092
QY	1230	-----TGAATTGAGGGAAGATGGAGCTCGTGTGCAATGACCAACGCCATC	1272
Db	1093	TTCTGCGAGTGGCAGCTCGAATTGATGGGAAGATGGAGCTCGTGTGCAATGACCAACGCCATC	1152
QY	1273	TTACTGACTGCTGTGATATATCTTCTGGAAGATGAGCCCTGTGACACTTTGGGATTT	1332
Db	1153	TTACTGACTGCTGTGATATATCTTCTGGAAGATGAGCCCTGTGACACTTTGGGATTT	1212
QY	1333	GCACCCATTAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGGT	1392
Db	1213	GCACCCATTAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGGT	1272
QY	1393	CTTGGAGTGCCTTACCATTTGCCATGAGTCTGGACACAACTTTGGCATGTATCATGAT	1452
Db	1273	CTTGGAGTGCCTTACCATTTGCCATGAGTCTGGACACAACTTTGGCATGTATCATGAT	1332
QY	1453	GGAGAAGGGAACAATGTGTAAAGATCCGAGGGCAACATCATGTCCCTTACATTTGGCAGGA	1512
Db	1333	GGAGAAGGGAACAATGTGTAAAGATCCGAGGGCAACATCATGTCCCTTACATTTGGCAGGA	1392
QY	1513	CGCAATGGAGTCTTCTCTGCTGATCAGCAAGCCCTGTGAAGCAATACAGATATCTCT	1632
Db	1393	CGCAATGGAGTCTTCTCTGCTGATCAGCAAGCCCTGTGAAGCAATACAGATATCTCT	1452
QY	1573	ACCGCTCAAGCTATCTGCTGATCAGCAAGCCCTGTGAAGCAATACAGATATCTCT	1632
Db	1453	ACCGCTCAAGCTATCTGCTGATCAGCAAGCCCTGTGAAGCAATACAGATATCTCT	1512
QY	1633	GAGAAATCCAGGAGAAATATATGATGCAAAACACACAGTGAAGTGGCAGTTCCGAGAG	1692
Db	1513	GAGAAATCCAGGAGAAATATATGATGCAAAACACACAGTGAAGTGGCAGTTCCGAGAG	1572
QY	1693	AAAGCCAGCTCTGCATCTGCATCTTAAAGAGACATCTGTAAGCCCTGTGTCGCAAT	1752
Db	1573	AAAGCCAGCTCTGCATCTGCATCTTAAAGAGACATCTGTAAGCCCTGTGTCGCAAT	1632
QY	1753	CGTATTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGCAACAATTTGTGGG	1812
Db	1633	CGTATTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGCAACAATTTGTGGG	1692
QY	1813	CATGACATGTGTGCGGGGAGACAGTGTGTGAATATGTTGATGAAGGCCCAAGCCC	1872
Db	1693	CATGACATGTGTGCGGGGAGACAGTGTGTGAATATGTTGATGAAGGCCCAAGCCC	1752
QY	1873	ACCCATGGCCACTGTGCGACTGGTCTTCTGTCCTCCATGCTCCAGGACCTCGGAGGG	1932
Db	1753	ACCCATGGCCACTGTGCGACTGGTCTTCTGTCCTCCATGCTCCAGGACCTCGGAGGG	1812
QY	1933	GGAGTATCTCATAGAGTCCCTCTGCACCAACCCCAAGCCATCGCATGGAGGGAAGTTC	1992
Db	1813	GGAGTATCTCATAGAGTCCCTCTGCACCAACCCCAAGCCATCGCATGGAGGGAAGTTC	1872
QY	1993	TGTGAGGCTCCACTCGCATCTGAGACTCTGCAACAGTCAGAAATGTCCCGGACAGT	2052
Db	1873	TGTGAGGCTCCACTCGCATCTGAGACTCTGCAACAGTCAGAAATGTCCCGGACAGT	1932
QY	2053	GTGTGACTTCCTGTGCTCAGTGTGCGGAGCAACAGCAGACGATTCAGAGGGCGGCAC	2112
Db	1933	GTGTGACTTCCTGTGCTCAGTGTGCGGAGCAACAGCAGACGATTCAGAGGGCGGCAC	1992
QY	2113	TACAAGTGAAGCCCTTACATCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATC	2172
Db	1993	TACAAGTGAAGCCCTTACATCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATC	2052
QY	2173	GCAGAAGATTTGATTTCTTCTTTTGTCAAATAAAGTCAAAGATGGGACTCCATGTC	2232
Db	2053	GCAGAAGATTTGATTTCTTCTTTTGTCAAATAAAGTCAAAGATGGGACTCCATGTC	2112
QY	2233	TCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAAT	2292
Db	2113	TCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAAT	2172
QY	2293	GTCTTGTGATCTGATGCTGTTGAAGACGCTCTGTGGGTTGTAAACGGGAATAAATCAAGC	2352
Db	2173	GTCTTGTGATCTGATGCTGTTGAAGACGCTCTGTGGGTTGTAAACGGGAATAAATCAAGC	2232
QY	2353	TGCAGATTCACAGGGGTCTCTACACCAAGCAGCACCACCAACCAACCACTATATCAGATG	2412
Db	2233	TGCAGATTCACAGGGGTCTCTACACCAAGCAGCACCACCAACCAACCACTATATCAGATG	2292
QY	2413	GTACCATTTCTTCTGGAGCCCGGAGTATCCGCATCTATGAATGAACGCTCTTACCTCC	2472
Db	2293	GTACCATTTCTTCTGGAGCCCGGAGTATCCGCATCTATGAATGAACGCTCTTACCTCC	2352
QY	2473	TACATTTCTGTGCGCAATGCGCTCTGAAAGGTACTACCTGAATGGGCACTGGACCGTGAC	2532
Db	2353	TACATTTCTGTGCGCAATGCGCTCTGAAAGGTACTACCTGAATGGGCACTGGACCGTGAC	2412
QY	2533	TGGCCCGCGCGTACAAATTTTTCGGGCACTACTTTTCGACTACAGACGCTCTTATATGAG	2592
Db	2413	TGGCCCGCGCGTACAAATTTTTCGGGCACTACTTTTCGACTACAGACGCTCTTATATGAG	2472
QY	2593	CCCGAAGACTTAATCGCTACTTGGACCAACCAACGAGACACTGATTTGTGGAGCTGCTGTT	2652
Db	2473	CCCGAAGACTTAATCGCTACTTGGACCAACCAACGAGACACTGATTTGTGGAGCTGCTGTT	2532
QY	2653	CAGGGAAGGACCCCGGTGTTTCCCTGGGAATACTCCATGCTCGTGGGACCGGAGAG	2712
Db	2533	CAGGGAAGGACCCCGGTGTTTCCCTGGGAATACTCCATGCTCGTGGGACCGGAGAG	2592
QY	2713	CAGCCCCCTGCCAGCCAGCTACACTTGGGCAATCTGTCGCTCTGAGTGTCCGTGTC	2772
Db	2593	CAGCCCCCTGCCAGCCAGCTACACTTGGGCAATCTGTCGCTCTGAGTGTCCGTGTC	2712
QY	2773	TGCGAGGGGG 2783	
Db	2653	TGCGAGGGGG 2663	
RESULT 4			
AX319852			
LOCUS	AX319852	3675 bp	DNA linear PAT 14-DEC-2001
DEFINITION	Sequence 16 from Patent WO0183782.		
ACCESSION	AX319852		
VERSION	AX319852.1	GI:17901442	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Plowman, G.D., Whyte, D., Sudarsanam, S., Manning, G., Caenepael, S. and Payne, V.		
TITLE	Novel proteases		
JOURNAL	Patent: WO 0183782-A 16 08-NOV-2001;		
FEATURES	Location/Qualifiers		
source	1..3675		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	874 a 1009 c 1019 g	773 t	
ORIGIN			
Query Match	64.9%;	Score 1877.8;	DB 6; Length 3675;
Best Local Similarity	84.4%;	Pred. No. 0;	
Matches 2289;	Conservative 0;	Mismatches 177;	Indels 245; Gaps 6;

QY 202 GCAGCACCTGGTGGCCATGGGACCCGACGCGGAGCGCTGGAGCCCGAGCGTCCC 261
Db 69 SCAGGCACTGGTGGCCATGGGACCCGACGCGGAGCGCTGGAGCCCGAGCGTCCC 128
QY 262 GCGTCTCTCTCACCGCGGAGCGCGCGGCTGGATGGAAAGGCGGAGCATGGATGAA 321
Db 129 GCGTCTCTCTCACCGCGGAGCGCGCGGCTGGATGGAAAGGCGG----- 175
QY 322 GCTGGAACCATCGTCTCAGCAAACTACACAGAAACAGAAACACAAACACTGCATGTT 381
Db 176 ----- 175
QY 382 CTCACCTCAATATGACCTGGTCTCTGCTACGAGGTTGACCAAGGCGGATTAAGTGTCC 441
Db 176 -----AATATGACCTGGTCTCTGCTACGAGGTTGACCAAGGCGGATTAAGTGTCC 228
QY 442 CATGAAATCATACCATCAGCGGCGGAGAGAGAGTGGCGTGTTCGAGGTTGAGTCT 501
Db 229 CATGAAATCATACCATCAGCGGCGGAGAGAGAGTGGCGTGTTCGAGGTTGAGTCT 288
QY 502 CTTACCTTCGGGTGAAGGCGGCGGAGAGAGTTCACATGGATCTGAGGACTTCCAGC 561
Db 289 CTTACCTTCGGGTGAAGGCGGCGGAGAGAGTTCACATGGATCTGAGGACTTCCAGC 348
QY 562 AGCCTAGTGGCTCTGGCTTTATTTGTCAGAGCGTTGGGAAAGACAGGCACTAAGTCTGTG 621
Db 349 AGCCTAGTGGCTCTGGCTTTATTTGTCAGAGCGTTGGGAAAGACAGGCACTAAGTCTGTG 408
QY 622 CAGACTTTACGCGCAGAGAGACTCTGTTCTATCAAGGCTCTTTGGGATCAACAGAAAC 681
Db 409 CAGACTTTACGCGCAGAGAGACTCTGTTCTATCAAGGCTCTTTGGGATCAACAGAAAC 468
QY 682 TCGCCATCGCATGAGGGAAGTTCTGTGAGGCGTCCACTCGCACTCTGAAGCTCTGCAAC 741
Db 469 TCTCTCA-----GTGGCCCTTTCAACCTGCGCAAGGCTGTGCA 504
QY 742 AGTCAGAAATGTCCCGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCCGAGCAAC 801
Db 505 GGCATGATACGAAACAGAGAGGAGATTAATCTCTAAGGCCACTTCTCTCACACCTCTCA 564
QY 802 AGCAGACGATTTAGAGGGGCGGACTACAGTGG--AAGCCTTACACTCAGTAGAGCGG 859
Db 565 TGGAAACTCGGAGAGCTGCCAAGGCACTGCCATCCCACTACAGAGATCC 624
QY 860 ACTTATGCAAACTCTACTGATCGCAGAGGATTTGATTTCTTTCTTTTGTCAATA 919
Db 625 ACAGAGCCCATGCTCTCTGGGCGAGTGAAGTCTCTGTGACCTC-----AAGG 672
QY 920 AAGTCAGAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATAT 979
Db 673 ACATGGGAGCTGGCACATCAACCCCTGCACAGCAGCGACCTTGGCTGGGACTGGCCAAA 732
QY 980 GTGAGCTCAGTGTGGTGTCCATCTGCG--GCACATGCCCCAGGCTCCCAAGGAGACCTC 1038
Db 733 AAGCAGCATTTCTGTGGAAGACCAAGAAATATGCTCCCGAGCTCCCAAGGAGACCTC 792
QY 1039 TTTCACTTTGCGCAGATGATATAAGTCTTGCTTACGGCATAAAGCGCTCTCTTCTGAGGTCC 1098
Db 793 TTTCACTTTGCGCAGATGATATAAGTCTTGCTTACGGCATAAAGCGCTCTCTTCTGAGGTCC 852
QY 1099 CATTAGAAATGAAGAACTGAACGTTGGAGACCTTGGTGGTGTGTCACAAAAGATGATGCAA 1158
Db 853 CATTAGAAATGAAGAACTGAACGTTGGAGACCTTGGTGGTGTGTCACAAAAGATGATGCAA 912
QY 1159 AACATGGCCATGAAATATCACCACTACGCTCAGGATCTCAGCATGATGATCTGCT 1218
Db 913 AACATGGCCATGAAATATCACCACTACGCTCAGGATCTCAGCATGATGATCTGCT 972
QY 1219 TTTATTCAAAGA----- 1229
Db 973 TTTATTCAAAGATGGAACAATAGGAGGAACATCAACATTCGAATTTGATAGTCTGATCTT 1032

QY 1230 ----- 1229
Db 1033 CTAGAGATGAACAGCCAGGACTCGTGATAAGTCAACACGAGACACACACCTTAAGTAGC 1092
QY 1230 -----TGGATTGATGGGGAAAAGATGGGACTCGTCATGACACACCCCATC 1272
Db 1093 TTTCCCAAGTGGCAGTCTGGATTGATGGGAAAAGATGGGACTCGTCATGACACCCCATC 1152
QY 1273 TTTACTGACTGCTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGACACATTTGGGATTT 1332
Db 1153 TTTACTGACTGCTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGACACATTTGGGATTT 1212
QY 1333 GCACCCATAAGTGGAAATGTGTAGTAAATATCCAGCTGCACAGATTAATGAAGATACAGGT 1392
Db 1213 GCACCCATAAGTGGAAATGTGTAGTAAATATCCAGCTGCACAGATTAATGAAGATACAGGT 1272
QY 1393 CTTGAGCTGGGCTTCAACCATTTGCCATGTGAGTCTGGACACAACTTTGGCATGATTCATGAT 1452
Db 1273 CTTGAGCTGGGCTTCAACCATTTGCCATGTGAGTCTGGACACAACTTTGGCATGATTCATGAT 1332
QY 1453 GGAGAGGGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCTCCCTACATTTGGCAGGA 1512
Db 1333 GGAGAGGGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCTCCCTACATTTGGCAGGA 1392
QY 1513 CGCAATGGAGTCTTCTCTGTGTCACCTGTCAGCCCGCAGTATCTACACAAATTTCTAAGC 1572
Db 1393 CGCAATGGAGTCTTCTCTGTGTCACCTGTCAGCCCGCAGTATCTACACAAATTTCTAAGC 1452
QY 1573 ACCGCTCAGCTATCTGCTGATCAGCCAAAGCCTGTCAAGGAATACAGATATCCT 1632
Db 1453 ACCGCTCAGCTATCTGCTGATCAGCCAAAGCCTGTCAAGGAATACAGATATCCT 1512
QY 1633 GAGAAATTCGCAAGAGAAATTAATGATGCAAAACACACAGTGCAGTGGCAGTTCGAGAG 1692
Db 1513 GAGAAATTCGCAAGAGAAATTAATGATGCAAAACACACAGTGCAGTGGCAGTTCGAGAG 1572
QY 1693 AAAGCCAGCTCTGATGCTGCACTTTAAAGGACATCTGTAAAGCCCTGTGTGTCCT 1752
Db 1573 AAAGCCAGCTCTGATGCTGCACTTTAAAGGACATCTGTAAAGCCCTGTGTGTCCT 1632
QY 1753 CGTATTGGAGGAAATGTGAGCTAAATTTATGCCAGCAGCAGAGAGGACAAATTTGTGG 1812
Db 1633 CGTATTGGAGGAAATGTGAGCTAAATTTATGCCAGCAGCAGAGAGGACAAATTTGTGG 1692
QY 1813 CATGACATGTGTGTCGGGAGGACAGTGTGTGAATATGTTGATGAAGGCCCAAGGCC 1872
Db 1693 CATGACATGTGTGTCGGGAGGACAGTGTGTGAATATGTTGATGAAGGCCCAAGGCC 1752
QY 1873 ACCCATGGCCACTGTGCTGGACTGCTTCTTGGTCCCATGCTCCAGGACCTGGCGAGGG 1932
Db 1753 ACCCATGGCCACTGTGCTGGACTGCTTCTTGGTCCCATGCTCCAGGACCTGGCGAGGG 1812
QY 1933 GGAGTATCTCATAGAGTGCCTCTGCACCAACCCCAAGCCATCGCATGGAGGAGTTTC 1992
Db 1813 GGAGTATCTCATAGAGTGCCTCTGCACCAACCCCAAGCCATCGCATGGAGGAGTTTC 1872
QY 1993 TGTGAGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTTCAGAAATGTCTCCCGGACAGT 2052
Db 1873 TGTGAGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTTCAGAAATGTCTCCCGGACAGT 1932
QY 2053 GTTGACTTCGCTGCTCAGTGTGCGGAGCAACAGCAGACGATTCAGAGGGCGGAC 2112
Db 1933 GTTGACTTCGCTGCTCAGTGTGCGGAGCAACAGCAGACGATTCAGAGGGCGGAC 1992
QY 2113 TACAAAGTGAAGCCCTTACACTCAAGTAGAAGTACAGGACTTATGCAAACTCTACTGTATC 2172
Db 1993 TACAAAGTGAAGCCCTTACACTCAAGTAGAAGTACAGGACTTATGCAAACTCTACTGTATC 2052
QY 2173 GCAGAGGATTTGATTTCTTTCTTTTGTGTCAAAATAAGATGAGGACTCCATGTC 2232
Db 2053 GCAGAGGATTTGATTTCTTTCTTTTGTGTCAAAATAAGATGAGGACTCCATGTC 2112
QY 2233 TCGGAGGATAGCCGCTTAATGTTTGTATAGATGGGATATGTGAGAGAGTGGATGTGACAA 2292

Db	2113	TCGGAGGATAGCGGTAAATGTTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAT	2172
Qy	2293	GTCTCTTGGATCTGATCCTGTGTGAAGACGCTGTGTGGGGTGTGTAAACGGGAATAAATCAGCC	2352
Db	2173	GTCTCTTGGATCTGATGCTGTGTGAAGACGCTGTGTGGGGTGTGTAAACGGGAATAAATCAGCC	2232
Qy	2353	TGCACGATTACAGAGGGTCTCTACACAAAGCACACACACCAACCAAGTATTATCACATG	2412
Db	2233	TGCACGATTACAGAGGGTCTCTACACAAAGCACACACACCAACCAAGTATTATCACATG	2292
Qy	2413	GTCAACCAATTCCTCTCTGGAGCCCGAGTATCCGCATCTATGAATGAACGTCTCTACCTCC	2472
Db	2293	GTCAACCAATTCCTCTCTGGAGCCCGAGTATCCGCATCTATGAATGAACGTCTCTACCTCC	2352
Qy	2473	TACATTTCTCTGGCAATGTCCTCAGAAAGTACTACTGAAATGGGCACTGGACCGTGGAC	2532
Db	2353	TACATTTCTCTGGCAATGTCCTCAGAAAGTACTACTGAAATGGGCACTGGACCGTGGAC	2412
Qy	2533	TGGCCCGCCGGTACAAATTTTGGGCACACTCTTCGACTACAGCGGTCCTATATGAG	2592
Db	2413	TGGCCCGCCGGTACAAATTTTGGGCACACTCTTCGACTACAGCGGTCCTATATGAG	2472
Qy	2593	CCCAGAGAACTTAATCGCTACTTGGACCAACCAACGAGACACTGATGTGTGGAGTCTGTTT	2652
Db	2473	CCCAGAGAACTTAATCGCTACTTGGACCAACCAACGAGACACTGATGTGTGGAGTCTGTTT	2532
Qy	2653	CAGGGAAGGAACCCGGGTGTTCCTGGGAATCTCCATGCTCGCTTGGGGACCGAGAG	2712
Db	2533	CAGGGAAGGAACCCGGGTGTTCCTGGGAATCTCCATGCTCGCTTGGGGACCGAGAG	2592
Qy	2713	CAGCCCTCTGCCAGCCAGCTACATTTGGGCCATCGTGGCTCTGAGTCTCCGTGTCC	2772
Db	2593	CAGCCCTCTGCCAGCCAGCTACATTTGGGCCATCGTGGCTCTGAGTCTCCGTGTCC	2652
Qy	2773	TGCGGAGGGG	2783
Db	2853	TGCGGAGGGG	2663

RESULT 5	BD170080	LOCUS	BD170080	linear	5610 bp	DNA	PAT 17-JAN-2003
DEFINITION	Novel ADAMTS family polypeptide and gene encoding the same.	ACCESSION	BD170080				
VERSION	BD170080.1	GI	27875892				
KEYWORDS	WO 0231163-A/2.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	1. (bases 1 to 5610) Obara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and Kanazaki,K.						
TITLE	Novel ADAMTS family polypeptide and gene encoding the same						
JOURNAL	Patent: WO 0231163-A 2 18-APR-2002; KAZUSA DNA RESEARCH INSTITUTE,MITSUBISHI PHARMA CORP.OSAMU OHARA, TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI MURAKAMI,SHINICHIRO YASUDA, KOJI KANZAKI						
COMMENT	OS Homo sapiens (human) PS WO 0231163-A/2 PD 18-APR-2002 PF 11-OCT-2001 WO 2001JP008913 PR 11-OCT-2000 JP 00P 311309,02-APR-2001 JP 01P 102905 PI OSAMU OHARA,TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI PI MURAKAMI, PI SHINICHIRO YASUDA,KOJI KANZAKI PC C12N15/57,C12N9/64,C12N5/00,C12N1/19,C12N1/21,C12N1/15,C07K16/ PC 40,C13Q1/58, PC G01N33/50 CC Novel ADAMTS family polypeptide and gene encoding the same FH Key Location/Qualifiers FT CDS (770) .(4444).						

FEATURES	source	Location/Qualifiers
BASE COUNT	1233 a	1638 c 1537 g 1202 t
ORIGIN		
Query Match	64.9%;	Score 1877.8; DB 6; Length 5610;
Best Local Similarity	84.4%;	Pred. No. 0;
Matches 2289; Conservative	0;	Mismatches 177; Indels 245; Gaps 6;
QY	202	GCAGGCACTTCGCTGCGCCATGGGACCCCGCAGCGCAGCGCTCTGGAGCCGAGCGTCC 261
DB	838	GCAGGCACTTCGCTGCGCCATGGGACCCCGCAGCGCAGCGCTCTGGAGCCGAGCGTCC 897
QY	262	GGGTCTCTTCACCCGCGAGCGCGCGGCTGGATGGAAAGGCGGACATGGATGAA 321
DB	898	GGGTCTCTTCACCCGCGAGCGCGCGGCTGGATGGAAAGGCGG----- 944
QY	322	GCTGGAAACCATCGTTCTCAGCAAACTAAACACAGGAACAGAAAAACCAAACTGCATGT 381
DB	945	----- 944
QY	382	CTCACTCAATATGACCTGGTCTCTGCTACGAGTTGACCAAGGGCGGATACGTGTCC 441
DB	945	-----AATATGACCTGGTCTCTGCTACGAGTTGACCAAGGGCGGATACGTGTCC 997
QY	442	CATGAATCATGACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGGTTGAGTCT 501
DB	998	CATGAATCATGACCAATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGGTTGAGTCT 1057
QY	502	CTTCACCTTCGGCTGAAAGCCCCCAGCACGACTTCCACATGATCTGAGGACTTCCAGC 561
DB	1058	CTTCACCTTCGGCTGAAAGCCCCCAGCACGACTTCCACATGATCTGAGGACTTCCAGC 1117
QY	562	AGCTAGTGCTCTCGCTTTATGTGCACAGCTTGGGAAGACAGCACTAAGTCTGTG 621
DB	1118	AGCTAGTGCTCTCGCTTTATGTGCACAGCTTGGGAAGACAGCACTAAGTCTGTG 1177
QY	622	CAGACTTTACCGCCAGAGGACTTCTGTGTTTCTATCAAGGCTCTTTGGCATCACACAGAAAC 681
DB	1178	CAGACTTTACCGCCAGAGGACTTCTGTGTTTCTATCAAGGCTCTTTGGCATCACACAGAAAC 1237
QY	682	TCGCCATCGCATGAGGAGGAAATTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAAC 741
DB	1238	TCCTCA-----GTGGCCCTTTCAACCTGCCAAGCGTTGTCA 1273
QY	742	AGTCAGAAATGCTCCCGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCGGAGCAAC 801
DB	1274	GGCATGATACGAACAGAGAGCGAGATTCTTCTTAAGGCCACTTCTTCCACACCTCTCA 1333
QY	802	AGCAGACGATTCAGAGGCGGCACTACAAGTGG--AAGCCCTTACACTCAAGTAGAAGCCG 859
DB	1334	TGGAAACTCGGCAGAGCTGCCCCAAGGAGCTCGCCATCCCAAGCTACTGTACAAGAGATCC 1393
QY	860	ACTTATGCAAACTCTACTGATCCGAGAGGATTTGATTTCTTCTTTCTTTGTCAAATA 919
DB	1394	ACAGAGCCCCATGCTCTCTGGGGCCAGTGAAGTCTCTGTGACCTC-----AAGG 1441
QY	920	AAGTCAAAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGTGGGATAT 979
DB	1442	ACATGGGAGCTGGCACATCAACCCCTGCACAGCAGCACTTCGCCCTGGGACTGCCACAA 1501
QY	980	GTGAGCTCAGTGTGGTCTCCACATCTCG--GCACATGCCCCAGCTCTCCAGGAGGACCTC 1038
DB	1502	AAGCAGCATTTCTGTGGAAGACGGAGAAATACATGCCCCAGCTCTCCAAAGGAAGACTC 1561
QY	1039	TTCACTTTCAGATGAGTATAGTCTTGTCTTACGGCATAGCGCTCTCTCTTGAGGTCC 1098
DB	1562	TTCACTTTCAGATGAGTATAGTCTTGTCTTACGGCATAGCGCTCTCTCTTGAGGTCC 1621
QY	1099	CATAGAAATGAAGAACTGAACGCTGGAGACCTTGGTGTGTCGACAAAAAGTATGCAC 1158

[illegible]

D	b	2702	GTTGACTTCCTCGTGTCTCAGTGCCGACCAACAAGCAGACGATTCACAGAGCGGGCCGCAC	2761
Q	y	2113	TACAAGTGGGAAGCCTTTACACTCAAGTAGAAGATCAGAGACTTATGCCAACTCTTACTGTATC	2172
D	b	2762	TACAAGTGGGAAGCCTTTACACTCAAGTAGAAGATCAGAGACTTATGCCAACTCTTACTGTATC	2821
Q	y	2173	GCAGAGGATTGATTTCTTCTTTCTTTGTGTAAGTCAAAGTCAAAGTGGGACTTCCATGC	2232
D	b	2822	GCAGAAGGATTGATTTCTTCTTTCTTTGTGTAAGTCAAAGTCAAAGTGGGACTTCCATGC	2881
Q	y	2233	TCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTCAGAGAGTTGGATGTGCAAT	2292
D	b	2882	TCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTCAGAGAGTTGGATGTGCAAT	2941
Q	y	2293	GTCTTGGATCTCATCTGTTGAAGACGCTCTGTTGGGTGTGTAACGGGAATAACTCAGCC	2352
D	b	2942	GTCTTGGATCTCATCTGTTGAAGACGCTCTGTTGGGTGTGTAACGGGAATAACTCAGCC	3001
Q	y	2353	TGCAGGATTACAGAGGGTCTCTACACCACACACACACACACACACACACACACACATG	2412
D	b	3002	TGCAGGATTACAGAGGGTCTCTACACCACACACACACACACACACACACACACACATG	3061
Q	y	2413	GTCAACATTCTTCTGAGCCCGGAGTATCCGCATCTATGAAATGAACGCTCTACCTCC	2472
D	b	3062	GTCAACATTCTTCTGAGCCCGGAGTATCCGCATCTATGAAATGAACGCTCTACCTCC	3121
Q	y	2473	TACATTCTGTGGCAATGCCCCACAGAGTACTACCTGAAATGAAAGTCTTACCTCC	2532
D	b	3122	TACATTCTGTGGCAATGCCCCACAGAGTACTACCTGAAATGAAAGTCTTACCTCC	3181
Q	y	2533	TGGCCCGCCGGTACAAAATTTTCGGGCACCTACTTTTCCACTACAGACGCTCTATAATGAG	2592
D	b	3182	TGGCCCGCCGGTACAAAATTTTCGGGCACCTACTTTTCCACTACAGACGCTCTATAATGAG	3241
Q	y	2593	CCGAGAACCTTAATCGCTACTTGGACACCAACCAACGAGACACTGATTTGGAGCTGCTGTT	2652
D	b	3242	CCGAGAACCTTAATCGCTACTTGGACACCAACCAACGAGACACTGATTTGGAGCTGCTGTT	3301
Q	y	2653	CAGGGAAGAACCCGGGTGTGCTGGGAATCTCCAATGCTCGCTTGGGGACCGGAGAAG	2712
D	b	3302	CAGGGAAGAACCCGGGTGTGCTGGGAATCTCCAATGCTCGCTTGGGGACCGGAGAAG	3361
Q	y	2713	CAGCCCCTCCAGACCCAGCTACACTTGGGCCATCGTGGCTGTGATGCTCGTGTCC	2772
D	b	3362	CAGCCCCTCCAGACCCAGCTACACTTGGGCCATCGTGGCTGTGATGCTCGTGTCC	3421
Q	y	2773	TGGGAGGGGG 2783	
D	b	3422	TGGGAGGGGG 3432	
RESULT 6				
BD170083				
LOCUS				
DEFINITION				
Accession				
Version				
Keywords				
Source				
Organism				
Reference				
Authors				
Title				
Journal				
Comment				

920 AAGTCAAGAGATGGACTCCATGCTCGAGGATAGCCGTAATGTTTGTATAGATGGGATAT 979
4724 ACATGGAGCTGGACATCAACCCCTGCACAGCAGGACCTTCGCCCTGGGACTGCCACAA 4783
980 GTGAGCTCACTGTGGTGTCCACATCTGC-GCAGATGCCGCCAGCTCCCAAGGAGACCTC 1038
4784 AAGCAGCATTTCTGTGGAGAGCGAAGAAATACATGCCCGAGCTCCCAAGGAGACCTC 4843
1039 TTCTATCTTCGAGATGAGTATAGTCTTCTTACGGCATAGCGCTCTCTTCTGAGGTCC 1098
4844 TTCTATCTTCGAGATGAGTATAGTCTTCTTACGGCATAGCGCTCTCTTCTGAGGTCC 4903
1099 CATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGTGGTGGTGGTGGTGGTGGTGGT 1158
4904 CATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGTGGTGGTGGTGGTGGTGGTGGT 4963
1159 AACCATGGCCATGAAATATACACCTACCTGCTTACAGTATCTCAACATGCTATCTGCT 1218
4964 AACCATGGCCATGAAATATACACCTACCTGCTTACAGTATCTCAACATGCTATCTGCT 5023
1219 TTATTCGAAG- 1229
5024 TTATTCGAAGATGGAACATAGGAGGAAACATCAACATTCGAATTTAGGTCTGATCTT 5083
1230 ----- 1229
5084 CTAGAGATGAACACCGAGGACTGGTATAGTCAACGACGACACACCTTAAAGTAGC 5143
1230 -----TGGATTGATGGGAAAGATGGGACTGGCTCGTATGACACGACCTC 1272
5144 TTCTGCCAGTGGCAGTCTGGGATTTGATGGGAAAGATGGGACTGGCTCGTATGACACGACCTC 5203
1273 TTCTGACTGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGACACTTTGGGATTT 1332
5204 TTCTGACTGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGACACTTTGGGATTT 5263
1333 GCACCATTAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGT 1392
5264 GCACCATTAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGT 5323
1393 CTTGGACTGGCTTCACCATTTGCCATGCTGAGTCTGACACAACTTTGGCATGATCATGAT 1452
5324 CTTGGACTGGCTTCACCATTTGCCATGCTGAGTCTGACACAACTTTGGCATGATCATGAT 5383
1453 GGAGAGGAGAACTGTGTAAAGTCCGAGGCGAAACATCATGTCCTTACATGTCGACGA 1512
5384 GGAGAGGAGAACTGTGTAAAGTCCGAGGCGAAACATCATGTCCTTACATGTCGACGA 5443
1513 CGCAATGAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
5444 CGCAATGAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5503
1573 ACCGCTCAAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
5504 ACCGCTCAAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5563
1633 GAGAAATTCGCGAGGAGATTTATGATGCAAAACACACAGTGCAGTGCAGTGCAGTGCAG 1692
5564 GAGAAATTCGCGAGGAGATTTATGATGCAAAACACACAGTGCAGTGCAGTGCAGTGCAG 5623
1693 AAAGCAGAGCTCTGATGCTGAGCTTTTAAAGAGGACATCTGTAAGCCCTGCTGCTGCTGCT 1752
5624 AAAGCAGAGCTCTGATGCTGAGCTTTTAAAGAGGACATCTGTAAGCCCTGCTGCTGCTGCT 5683
1753 CGTATTGGAAGAAATGTGAGACTAAATTTATGCGAGCAGCAGAGGAGGACAAATTTGTGG 1812
5684 CGTATTGGAAGAAATGTGAGACTAAATTTATGCGAGCAGCAGAGGAGGACAAATTTGTGG 5743
1813 CATGACATGTGTGCGGGAGGACAGTGTGTGAAATATGTTGATGTAAGGCGCCCAAGGCC 1872
5744 CATGACATGTGTGCGGGAGGACAGTGTGTGAAATATGTTGATGTAAGGCGCCCAAGGCC 5803
1873 ACCCATGGCCACTGCTGCGAGTGTCTTCTTGGTCCCATGCTCCAGGACCTGCGGAGGG 1932

PR 11-OCT-2000 JP OOP 311309.02-APR-2001 JP OIP 102905 PI
OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI FI
MURAKAMI,
PI SHINICHIRO YASUDA, KOJI KANZAKI
PC C12N15/57, C12N9/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/
PC 40, C12Q1/68,
PC G01N33/50
CC Description of Artificial Sequence: base sequence of the CC
plasmid
FH pFastBac1-HT-PJ01256 Location/Qualifiers
FT key 1..8435 /organism='Artificial Sequence'.
FT source 1..8435 /organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
BASE COUNT 2072 a 2174 c 2149 g 2040 t
ORIGIN
Query Match 64.9%; Score 1877.8; DB 6; Length 8435;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2289; Conservative 0; Mismatches 177; Indels 245; Gaps 6;
QY 202 GCAGGACCTGCTGGCCATGGGACCCGCGGCGCGCGCTGGGAGCCCGAGCGTCCC 261
DB 4120 GCAGGACCTGCTGGCCATGGGACCCGCGGCGCGCGCTGGGAGCCCGAGCGTCCC 4179
QY 262 GCGTCTCTCCACCCGCGGAGCGCGCGGCTGGATGGAAGAGCGGAGCATGGATGAA 321
DB 4180 GCGTCTCTCCACCCGCGGAGCGCGCGGCTGGATGGAAGAGCGG----- 4226
QY 322 GCTGGAACCATGCTTCTCAGCAAACTAACAGAGACAGAAACACCAACTGCTGTT 381
DB 4227 ----- 4226
QY 382 CTCACCTCAATATGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
DB 4227 -----AATATGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4279
QY 442 CATGAAATCATGACCATCAGCGGCGGAGAGAGCAGTGGCGGTGCTGCGAGGTGAGTCT 501
DB 4280 CATGAAATCATGACCATCAGCGGCGGAGAGAGCAGTGGCGGTGCTGCGAGGTGAGTCT 4339
QY 502 CTTACCTTGGCTGAAAGGCCCGGAGGAGTCCACATGATCTGAGGACTTCCAGC 561
DB 4340 CTTACCTTGGCTGAAAGGCCCGGAGGAGTCCACATGATCTGAGGACTTCCAGC 4399
QY 562 AGCTAGTGGCTCCTGGCTTTATGTCAGACGCTTGGGAAAGACAGGCACTAAGTCTGTG 621
DB 4400 AGCTAGTGGCTCCTGGCTTTATGTCAGACGCTTGGGAAAGACAGGCACTAAGTCTGTG 4459
QY 622 CAGACTTTACCGCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGATCAGACAGAAC 681
DB 4460 CAGACTTTACCGCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGATCAGACAGAAC 4519
QY 682 TCGCCATCGCATGGAGGAGTCTGTGAGGGCTTCCACTCGCACTCTGAGGCTCTGCAAC 741
DB 4520 TCCTCA-----GTGGCCCTTTCACCTCTGCGAGGCTTGTCA 4555
QY 742 AGTCAGAAATGTCCCGGAGAGTGTGACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
DB 4556 GGCATGATGACGACAGAGAGGAGGAGTACTTCTAAGGCCACTTCTCTCACACCTCTCA 4615
QY 802 AGCAGAGATTCAGAGGGCGGCACTACAAGTGG--AAGCCTTACACTCAAGTGAAGCGG 859
DB 4616 TGGAAACTCGGCAAGGAGTCTGTTTCTATCAAGGCTTTTGGATCAGACAGATCC 4675
QY 860 ACTTATCAAACTCTACTGATGACAGAGGATTTGATTTCTTTCTTTCTTTCTTTCTTTCT 919
DB 4676 ACAGAGCCCATGCTCTCTGGGCGCAGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 4723


```

Db 5804 ACCATGGCCACTGGTCGGACTGGTCTCTGTGTCCTCCAGGACCTGCGGAGG 5863
Qy 1933 GGAGTATCTCATAGGAGTGGCTCTGCGACCAACCCCAAGCCATCGCATGGAGGAGTTC 1992
Db 5864 GGAGTATCTCATAGGAGTGGCTCTGCGACCAACCCCAAGCCATCGCATGGAGGAGTTC 5923
Qy 1993 TGTGAGGGTCCCACTGCGACTCTGAGCTCTGCAACAGTGCAGAAATGTCCCGGGACAGT 2052
Db 5924 TGTGAGGGTCCCACTGCGACTCTGAGCTCTGCAACAGTGCAGAAATGTCCCGGGACAGT 5983
Qy 2053 GTTGACTTCCGTGCTGCTGAGTGTGCGAGCACAACAGCAGATTCAGAGGCGGAC 2112
Db 5984 GTTGACTTCCGTGCTGCTGAGTGTGCGAGCACAACAGCAGATTCAGAGGCGGAC 6043
Qy 2113 TACAAGTGAAGCCTTACACTCAAGTGAAGATCAGACATTAATGCAAACTCTACTGATC 2172
Db 6044 TACAAGTGAAGCCTTACACTCAAGTGAAGATCAGACATTAATGCAAACTCTACTGATC 6103
Qy 2173 GCAGAAGGATTGATTTCTTTCTTTCTGTCAAATAAAGTCAAAAGATGGGACTCCATGC 2232
Db 6104 GCAGAAGGATTGATTTCTTTCTTTCTGTCAAATAAAGTCAAAAGATGGGACTCCATGC 6163
Qy 2233 TCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGCAAT 2292
Db 6164 TCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGCAAT 6223
Qy 2293 GTCTTGGATGATGCTGTGAAGAGCTCTGCGGTGTGTAACGGGATTAATCAATG 2352
Db 6224 GTCTTGGATGATGCTGTGAAGAGCTCTGCGGTGTGTAACGGGATTAATCAATG 6283
Qy 2353 TGCACAGATTACAGGGTCTCTACACCAAGCACCACACACCAACCAACCAATTAATCAATG 2412
Db 6284 TGCACAGATTACAGGGTCTCTACACCAAGCACCACACACCAACCAACCAATTAATCAATG 6343
Qy 2413 GTCAACATTCCTTCTGAGAGCCCGAGTATCCGATCTATGAAATGAACGCTCTTACCTCC 2472
Db 6344 GTCAACATTCCTTCTGAGAGCCCGAGTATCCGATCTATGAAATGAACGCTCTTACCTCC 6403
Qy 2473 TACATTTCTGTGCGCATGCCCCCTCAGAGGTAATCTGATGAGTGGGACCGTGGAC 2532
Db 6404 TACATTTCTGTGCGCATGCCCCCTCAGAGGTAATCTGATGAGTGGGACCGTGGAC 6463
Qy 2533 TGGCCCGCCCGGTACAAATTTTCGGGCACTACTTTGCACTACAGACGCTCTTAATGAG 2592
Db 6464 TGGCCCGCCCGGTACAAATTTTCGGGCACTACTTTGCACTACAGACGCTCTTAATGAG 6523
Qy 2593 CCCGAGAACTTAATCGCTACTGAGACCAACCAACGAGACACTGATTTGGAGCTCTCTTT 2652
Db 6524 CCCGAGAACTTAATCGCTACTGAGACCAACCAACGAGACACTGATTTGGAGCTCTCTTT 6583
Qy 2653 CAGGGAAGAAACCCGGGTGTTGCTTGGGAATACTCCATGCTCGCTTGGGACCGAGAAG 2712
Db 6584 CAGGGAAGAAACCCGGGTGTTGCTTGGGAATACTCCATGCTCGCTTGGGACCGAGAAG 6643
Qy 2713 CAGCCCTCCCGACCCAGCTACACTTTGGGCAATGCTGCGCTCTGAGTGTCCGCTGCC 2772
Db 6644 CAGCCCTCCCGACCCAGCTACACTTTGGGCAATGCTGCGCTCTGAGTGTCCGCTGCC 6703
Qy 2773 TCGGAGAGGGG 2783
Db 6704 TCGGAGAGGGG 6714

```

RESULT 7

```

LOCUS BD170084 8505 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel ADAMTS family polypeptide and gene encoding the same.
ACCESSION BD170084
VERSION BD170084.1 GI:27875896
KEYWORDS WO 0231163-A/6.
SOURCE synthetic construct
ORGANISM

```

```

artificial sequences.
1 (bases 1 to 8505)
Ohara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and
Kanzaki,K.
Novel ADAMTS family polypeptide and gene encoding the same
Patent: WO 0231163-A 6 18-APR-2002;
KAZUSA DNA RESEARCH INSTITUTE,MITSUBISHI PHARMA CORP,OSAMU OHARA,
TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI MURAKAMI,SHINICHIRO
YASUDA,KOJI KANZAKI
OS Artificial Sequence
PN WO 0231163-A/6
PD 18-APR-2002 WO 2001JP008913
PF 11-OCT-2001 WO 2001JP008913
PR 11-OCT-2000 JP 00P 311309,02-APR-2001 JP 01P 102905 PI
OSAMU OHARA,TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI PI
MURAKAMI,
PI SHINICHIRO YASUDA,KOJI KANZAKI
PC C12N15/57,C12N9/64,C12N5/00,C12N1/19,C12N1/21,C12N1/15,C07K16/
PC 40,C12Q1/68,
PC G01N33/50
CC Description of Artificial Sequence:base sequence of the CC
plasmid
CC pFastBac1-MS/HT-PJ01256-2
FH Key Location/Qualifiers
FT source 1..8505
/organism='Artificial Sequence'
/Plasmid
FEATURES
source
1..8505
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 2086 a 2188 c 2164 g 2067 t
ORIGIN
Query Match 64.9%; Score 1877.8; DB 6; Length 8505;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2289; Conservative 0; Mismatches 177; Indels 245; Gaps 6;
Qy 202 GCAGGCACCTCGCTGCGCCATGGAGCCCGCAGCGGAGCGCCCTGGGAGCCCGAGCGTCC 261
Db 4190 GCAGGCACCTCGCTGCGCCATGGAGCCCGCAGCGGAGCGCCCTGGGAGCCCGAGCGTCC 4249
Qy 262 GCGTCTCTCCACCCCGGAGCGCGGCTGGATGGAAGGCGGACATGGATGAA 321
Db 4250 GCGTCTCTCCACCCCGGAGCGCGGCTGGATGGAAGGCGGCG- 4296
Qy 322 GCTGGAACCATCGTTCTCAGCAAACTAACACAGGAACCAAAACCACTGCTATGTT 381
Db 4297 ----- 4296
Qy 382 CTCACCTCAATATGACCTGGTCTCTGCTACGAGTTGACACAGGGCGGATACGTTGCC 441
Db 4297 -----AATATGACCTGGTCTCTGCTACGAGTTGACACAGGGCGGATACGTTGCC 4349
Qy 442 CATGAATCATGCACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCGGAGTTGAGTCT 501
Db 4350 CATGAATCATGCACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCGGAGTTGAGTCT 4409
Qy 502 CTTCACTCTCGCTGAAAGGCCCGCAGGCAGACTTCCACATGGAATCTGAGGACTTCCAGC 561
Db 4410 CTTCACTCTCGCTGAAAGGCCCGCAGGCAGACTTCCACATGGAATCTGAGGACTTCCAGC 4469
Qy 562 AGCCTAGTGGCTCTGGCTTTATGTGACAGCGTGGGAAAGACAGGACATTAAGTCTGTG 621
Db 4470 AGCCTAGTGGCTCTGGCTTTATGTGACAGCGTGGGAAAGACAGGACATTAAGTCTGTG 4529
Qy 622 CAGCTTTACCGCCAGAGGACTCTGTTCTATCAAGGCTCTTTGCGATCACAGAAAC 681
Db 4530 CAGCTTTACCGCCAGAGGACTCTGTTCTATCAAGGCTCTTTGCGATCACAGAAAC 4589
Qy 682 TCGCATCGCATGAGGGAAGTCTGTGTGAGGGGTCCACTCGCACTCTGAAGCTCTGCAAC 741
Db 4590 TCCTCA-----GTGCCCTTTCAACCTGCCAAGGCTTGTCAC 4625

```


QY	742	AGT CAGAAATGTCCTCCCGGGACAGGTGTTGAATCTCCGTCGTGCTCAGTGTCGCCGAGCACCAAC	801
Dd	4626	GGCATGATACGAACAGAAGAGGAGATTACTCTTAAGGCACCTTCTTACACACCTCTCA	4685
QY	802	AGCAGACGATTCAGAGGGCGGCACTACAAGTGG--AAGCCTTACACTCAAGTAGAAGCCG	859
Dd	4686	TGGAACCTCGCAGAGCTGCCAACAGGAGCTCGCCATCCCACGTACTGTACAAGATCC	4745
QY	860	ACTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTCTTCTTCTTCTTCTCAAATA	919
Dd	4746	ACAGAGCCCATGCTCTCTGGGGCAGTGAGGTCTCTGTGACCTC-----AAGG	4793
QY	920	AAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATAT	979
Dd	4794	ACATGGAGCTGGCAGACTCAACCCTCTGCACAGCAGCACCTTGCCTGGGACTGCCACAA	4853
QY	980	GAGAGCTCAGTGTGGTGTCCACATCTGC-GCACATGCCAGCCCTCCCAAGGAGACCTC	1038
Dd	4854	AAGCAGCATTTCTGTGGNAAGCGCAAGAAATACATATGCTCCCGCCCTCCCAAGGAGACCTC	4913
QY	1039	TTTATCTTCCAGATGAGTATAAGTCTTTGCTTACGGCATAGCCCTCTCTTCTGAGTCC	1098
Dd	4914	TTTATCTTCCAGATGAGTATAAGTCTTTGCTTACGGCATAGCCCTCTCTTCTGAGTCC	4973
QY	1099	CATAGAAATGAAGAACTGAACGTTGGAGACCTTGGTGTGGTGTGCGACAAAAGATGATCAA	1158
Dd	4974	CATAGAAATGAAGAACTGAACGTTGGAGACCTTGGTGTGGTGTGCGACAAAAGATGATCAA	5033
QY	1159	RACCATGGCATGAAAAATACACACCTPACGTGCTCAGCATCTCAACATGGTATCTGCT	1218
Dd	5034	AACCATGGCATGAAAAATACACACCTPACGTGCTCAGCATCTCAACATGGTATCTGCT	5093
QY	1219	TTATTTCAAAGA-----	1229
Dd	5094	TTATTTCAAAGTGAACAAATAGGAGGAAACATCMACATTGCAATTGTAGTCTGATCTT	5153
QY	1230	-----	1229
Dd	5154	CTAGAAATGAACACCGCAGGACTGTGTATAGTCAACCGAGCACCAACCTTAAGTAGC	5213
QY	1230	-----TGGAATTTGGGAAAGATGGGACTGTGTCATGACACGCCATC	1272
Dd	5214	TTCTGCCAGTGGCAGTCTGGATTTGATGGGAAAGATGGGACTGTGTCATGACACGCCATC	5273
QY	1273	TTACTGACTGGTCTGGATATATGTTCTTGGAAAGATAGGCCCTGTGACACTTTGGGATTT	1332
Dd	5274	TTACTGACTGGTCTGGATATATGTTCTTGGAAAGATAGGCCCTGTGACACTTTGGGATTT	5333
QY	1333	GNCCCATAGTGAATGTGTAGTAAATATCCAGCTCCAGCTCAACGATTAATGAAGATACAGST	1392
Dd	5334	GCACCATAGTGAATGTGTAGTAAATATCCAGCTCCAGCTCAACGATTAATGAAGATACAGST	5393
QY	1393	CTTGACTGGCTTCAACATGCCCCATGAGTCTGGACACAACTTTGGCATGATTCATGAT	1452
Dd	5394	CTTGACTGGCTTCAACATGCCCCATGAGTCTGGACACAACTTTGGCATGATTCATGAT	5453
QY	1453	GGHAAAGGGAACATGTTAAAAGTCCGAGGGCAACATCATGTCCCTTACATTTGGCAGGA	1512
Dd	5454	GGHAAAGGGAACATGTTAAAAGTCCGAGGGCAACATCATGTCCCTTACATTTGGCAGGA	5513
QY	1513	CGCAATGGAGTCTTCTCTGTCACCTCCAGCCGAGTATCTACACAAATTTCTTAAGC	1572
Dd	5514	CGCAATGGAGTCTTCTCTGTCACCTCCAGCCGAGTATCTACACAAATTTCTTAAGC	5573
QY	1573	ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGGAATACAAATATCTT	1632
Dd	5574	ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGGAATACAAATATCTT	5633
QY	1633	GAGAAATTCGAGAGAAATATATGATGMAACACACAGTGCMAAGTGGCAGTTCGGAGAG	1692
Dd	5634	GAGAAATTCGAGAGAAATATATGATGMAACACACAGTGCMAAGTGGCAGTTCGGAGAG	5693

Db	6774	TGCGAGGGGG	6784
<hr/>			
RESULT 8			
AR229456			
LOCUS AR229456 linear PAT 20-DEC-2002			
DEFINITION Sequence 3 from patent US 6448388.			
ACCESSION AR229456			
VERSION AR229456.1 GI:27268950			
KEYWORDS Unknown.			
SOURCE Unknown.			
ORGANISM Unclassified.			
REFERENCE 1 (bases 1 to 3675)			
AUTHORS Fiddle,C.J. and Hilbun,E.			
TITLE Human proteases and polynucleotides encoding the same			
JOURNAL Patent: US 6448388-A 3 10-SEP-2002;			
FEATURES Location/Qualifiers			
source 1..3675			
BASE COUNT 873 a 1008 c 1020 g 774 t			
ORIGIN /organism="unknown"			
Query Match 64.8%; Score 1874.6; DB 6; Length 3675;			
Best Local Similarity 84.4%; Pred. No. 0;			
Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;			
Qy	202	GCAGGACCTGTGGCCATGGGACCAGCGCAGCGCTGGAGCCCGAGGTCCTCC	261
Dd	69	GCAGGACCTGTGGCCATGGGACCAGCGCAGCGCTGGAGCCCGAGGTCCTCC	128
Qy	262	GCGTCTCTCCACCGCGAGCGCGGCTGGATGGAAGAAGCGCGCATGGATCAA	321
Dd	129	GCGTCTCTCCACCGCGAGCGCGGCTGGATGGAAGAAGCGCG-----	175
Qy	322	GCTGGAAACCATCGTTCTCAGCAACTAACA CAGGAACAGAAAACCAACTGCATGT	381
Dd	176	-----	175
Qy	382	CTCACTCAATATACCTGGTCTTGCTTAGAGGTTGACACAGGGCGATTACGTGCC	441
Dd	176	-----AATATACCTGGTCTTGCTTAGAGGTTGACACAGGGCGATTACGTGCC	228
Qy	442	CATGAATCATCACCATCAGCGCGGAGAGAGAGTGGCGTCTCGAGGTTGAGTCT	501
Dd	229	CATGAATCATCACCATCAGCGCGGAGAGAGAGTGGCGTCTCGAGGTTGAGTCT	288
Qy	502	CTTCACCTTCGGCTGAAAGCCCCAGCGACACTTCCACATGGATCTCAGGACTTCCAGC	561
Dd	289	CTTCACCTTCGGCTGAAAGCCCCAGCGACACTTCCACATGGATCTCAGGACTTCCAGC	348
Qy	562	AGCTAGTGTCTCGCTTATTTGTCAGAGCTTGGGAGACAGGCACTTAAGTCTGTG	621
Dd	349	AGCTAGTGTCTCGCTTATTTGTCAGAGCTTGGGAGACAGGCACTTAAGTCTGTG	408
Qy	622	CAGACTTTACGCGCAGAGGACTTCTGTTCTATCAAGGCTCTTGGCATCACAGAAAC	681
Dd	409	CAGACTTTACGCGCAGAGGACTTCTGTTCTATCAAGGCTCTTGGCATCACAGAAAC	468
Qy	682	TCGCCATCGCATGGAGGAGTTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAAC	741
Dd	469	TCCTCA-----GTGGCCCTTTCAACCTGCCAGGCTTGTC	504
Qy	742	AGTCAGAAATGTCCCGGGACAGTGTGTGACTTCGCTGCTGCTCAGTGTGCGAGCAAC	801
Dd	505	GGCATGATACGAACAGAGAGGAGCATTAATCTTAAAGGCCACTTCTTCAACCTCTCA	564
Qy	802	AGCAGACGATTTCAGAGGGCGGCACTCAAGTGG--AAGGCTTTACACTCAAGTAAAGCGG	859
Dd	565	TGGAATCTCGGAGAGCTGCCAGGAGGCTGCCATCCACAGTACTGTACAGAGATCC	624
Qy	860	ACTTATGCAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTTGTCAATA	919

Db	1893	1873	1753	1933	1813	1993	1873	2053	1933	2113	1993	2173	2053	2233	2113	2293	2173	2353	2233	2413	2293	2473	2353	2533	2413	2593	2473	2553	2713	2593	2773	2653	
Qy	CATGACATGTGGTCCCGGGAGGACAGTGTGTGAAATATGCTGATGAAGGCCCAAGCCC	ACCATGGCCACTGGTTCGGAGTGGTCTCTTGGTCCCACAGCTCTCCAGGAGG	ACCATGGCCACTGGTTCGGAGTGGTCTCTTGGTCCCACAGCTCTCCAGGAGG	GGAGTATCTCATAGAGTGGCTCTGCAACCAACCCCAAGGCATCGCATGGAGGGAAGTTC	GGAGTATCTCATAGAGTGGCTCTGCAACCAACCCCAAGGCATCGCATGGAGGGAAGTTC	TGTGAGGCTCCACTGGCACTCTGAGACTCTGACAGCTCTGACACAGTCTGAGAAATGTCCCGGGACAGT	TGTGAGGCTCCACTGGCACTCTGAGACTCTGACAGCTCTGACAGTCTGAGAAATGTCCCGGGACAGT	GTGACTTCCGTCGTCTGCTCAGTGTGCCAGACACAAACAGACAGGATTCAGAGGCGGCAC	GTGACTTCCGTCGTCTGCTCAGTGTGCCAGACACAAACAGACAGGATTCAGAGGCGGCAC	TACAAGTGGAGCCCTTACACTCAAGTAGAGATCAGGACTTATGCGAACTCTACTGTATC	TACAAGTGGAGCCCTTACACTCAAGTAGAGATCAGGACTTATGCGAACTCTACTGTATC	GCAGAGGAGTTGATTTCTCTCTTTCTTTGTCAAATAAAGTCAAAAGATGGGACTCCATGC	GCAGAGGAGTTGATTTCTCTCTTTCTTTGTCAAATAAAGTCAAAAGATGGGACTCCATGC	TCGGAGATAGCGGTAATGTTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAAT	TCGGAGATAGCGGTAATGTTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAAT	GTCTTTGGATCTGATCTCTGTGAAGACGCTCTGTGGGGTGTGTAACGGGAATAACTCAGCC	GTCTTTGGATCTGATCTCTGTGAAGACGCTCTGTGGGGTGTGTAACGGGAATAACTCAGCC	TGCAAGATTCACAGGGGTCTCTACACAAAGCACCAACACCAACAGTATTCACATG	TGCAAGATTCACAGGGGTCTCTACACAAAGCACCAACACCAACAGTATTCACATG	GTCAAGATTCCTTCTCGAGCCCGAGTATCCGGCATCTATGAAATGAACGCTCTACCTCC	GTCAAGATTCCTTCTCGAGCCCGAGTATCCGGCATCTATGAAATGAACGCTCTACCTCC	TACATTTCTGTGGCAATGCCCTCAGAGGTACTACTGAATGGGCACCTGGACCTGGAC	TACATTTCTGTGGCAATGCCCTCAGAGGTACTACTGAATGGGCACCTGGACCTGGAC	TGGCCCGCGGTACAAATTTTCGGGCACCTACTTTTCGACTACAGAGCGTCTTAAATGAG	TGGCCCGCGGTACAAATTTTCGGGCACCTACTTTTCGACTACAGAGCGTCTTAAATGAG	CCCGAGAACTTAATTCGCTACTGGACCAACCAACAGACACACTGATTTGGAGCTGTCTGTTT	CCCGAGAACTTAATTCGCTACTGGACCAACCAACAGACACACTGATTTGGAGCTGTCTGTTT	CAGGAGAGAACCCGGGTGTGCTGGGAATCTCCATGCTCTGGTGGGACCCGAGAAG	CAGGAGAGAACCCGGGTGTGCTGGGAATCTCCATGCTCTGGTGGGACCCGAGAAG	CAGCCCCCTGCCAGCCCCAGCTACACTTTGGGCCCATGTCGCTCTGAGTCTCCGTGTC	CAGCCCCCTGCCAGCCCCAGCTACACTTTGGGCCCATGTCGCTCTGAGTCTCCGTGTC	TGCGGAGGGGG 2783	TGCGGAGGGGG 2663
Db																																	

RESULT 9

BD170564

LOCUS

DEFINITION

ACCESSION

VERSION

BD170564

Novel protease.

BD170564

BD170564.1

GI:27876376

BD170564

linear

PAT 17-JAN-2003

RESULT 10
AR229457
LOCUS AR229457 4042 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 5 from patent US 6448388.
ACCESSION AR229457
VERSION AR229457.1 GI:27268951
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 4042)
Fridler-C.J. and Hilbun,B.
AUTHORS Human proteases and polynucleotides encoding the same
TITLE Patent: US 6448388-A 5 10-SEP-2002;
FEATURES
Location/Qualifiers
1..4042
/organism="unknown"
BASE COUNT 947 a 1117 c 1130 g 848 t
ORIGIN
Query Match 64.8%; Score 1874.6; DB 6; Length 4042;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;
QY 202 GCAGGACCTGCGTGGCCATGGACCCGAGCGGAGCGCCCTGGGAGCCGAGCGTCCC 261
DB 167 GCAGGACCTGCGTGGCCATGGACCCGAGCGGAGCGCCCTGGGAGCCGAGCGTCCC 226
QY 262 GCGTCTCTCCACCCGCGGAGCGCGGCTGGATGGAAGAGGGCGGACATGGATGAA 321
DB 227 GCGTCTCTCCACCCGCGGAGCGCGGCTGGATGGAAGAGGGCGG----- 273
QY 322 GCTGGAACCATGCTCTCAGCAAACTAACACAGGACGAAACCAACACTGCATGTT 391
DB 274 ----- 273
QY 382 CTCACCTCAATAGACCTGGTCTCTGCTACGAGTTGACACAGGGCGGATTAAGTGTCC 441
DB 274 -----AATATGACCTGGTCTCTGCTACGAGTTGACACAGGGCGGATTAAGTGTCC 326
QY 442 CATGAATCATGCAACATACGCGCGGAGAGAGAGTGCCTGCGGAGTTGAGTCT 501
DB 327 CATGAATCATGCAACATACGCGCGGAGAGAGAGTGCCTGCGGAGTTGAGTCT 386
QY 502 CTTACCTTCGGCTGAAGGCCCCAGGACGACTTCCACATGATCTGAGGACTTCCAGC 561
DB 387 CTTACCTTCGGCTGAAGGGCTCCGCGCAGCTTCCGCTGAGTCTGAGGACTTCCAGC 446
QY 562 AGCCTAGTGGCTCCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTG 621
DB 447 AGCCTAGTGGCTCCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTG 506
QY 622 CAGACTTACGGCCAGAGGACTTCTGTTCTATCAAGGCTTTCGATCAGGATCAGCAAGAAC 681
DB 507 CAGACTTACGGCCAGAGGACTTCTGTTCTATCAAGGCTTTCGATCAGGATCAGCAAGAAC 566
QY 682 TCGCCATCGCATGGAGGAGTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAAC 741
DB 567 TCCTCA-----GTGCCCCCTTCAACCTGCGAAGCTTGTCA 602
QY 742 AGTCAGAAATGTCGCCGAGCAGTGTGACTTCTGCTGCTGCTAGTGTGCGGAGCAAC 801
DB 603 GGCAATGATACGAACAGAGAGGCGAGTACTTCTTAAGGCGCACTTCTTCCACACCTCTCA 662
QY 802 AGCAGAGCATTCAGAGGCGGCACTACAGTGG--AAGCCCTTACACTCAGATGAGCCG 859
DB 663 TGGAAATCGGAGAGTGTGCCAAGGAGGCTTCGCCATCCCACTGATGTCAGAGATCC 722
QY 860 ACTTATGCAAACTCTACTGTATCGCAGAAAGGATTTGATTTCTTTCTTTTGTTCMAATA 919
DB 723 ACAGAGCCCCATGCTCCTGGGGCCAGTGAGGCTCTGCTGACCTC-----AAGG 770

QY 920 AAGTCARAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATATGATGGATAT 979
DB 771 ACATGGGAGCTGGGACATCAACCCCTGCACAGCAGGACCTTGGCTGGGACTGCCACNA 830
QY 980 GTGAGCTCAGTGTGTGTCCATCTGC-GCACATGCCCCAGGCTTCCCAAGGAAAGCTC 1038
DB 831 AAGCAGCATTTCTGTGGAAGAGCGCAAGAAATACATGCCCCAGCTTCCCAAGGAAAGCTC 890
QY 1039 TTCACTTTCGAGATGAGTATAGTCTTCTTACGGCATAGGCGCTCTCTCTGAGGTCC 1098
DB 891 TTCACTTTCGAGATGAGTATAGTCTTCTTACGGCATAGGCGCTCTCTCTGAGGTCC 950
QY 1099 CATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1158
DB 951 CATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1010
QY 1159 AACCATGGCCATGAAAAATATCACCACTTACGCTTACGATATCAATGATGATGATGATG 1218
DB 1011 AACCATGGCCATGAAAAATATCACCACTTACGCTTACGATATCAATGATGATGATGATG 1070
QY 1219 TTATTCAAGA----- 1229
DB 1071 TTATTCAAGAATGGAACAAATAGGAGGAAACATCAACATTGCAATTGTAGGTCTGATCTT 1130
QY 1230 ----- 1229
DB 1131 CTAGAAGATGAACAGCCAGGACTGTGTATAAGTCCACGACGACACACACCTTAAAGTAGC 1190
QY 1230 -----TGGATTGATGGGAAGATGGGACTGCTGTCATGACACACGCTC 1272
DB 1191 TTCTGCCAGTGCAGCTGTGGATTGATGGGAAGATGGGACTGCTGATGACACGCTC 1250
QY 1273 TTACTGACTGCTGTGATATATGTTCTTGGAAAGATAGCCCTGTGACACTTTGGGATTT 1332
DB 1251 TTACTGACTGCTGTGATATATGTTCTTGGAAAGATAGCCCTGTGACACTTTGGGATTT 1310
QY 1333 GCACCCATAAGTGGATGTAGTAAATATCCAGCTGCACGATTAAGTAAAGATACAGGT 1392
DB 1311 GCACCCATAAGTGGATGTAGTAAATATCCAGCTGCACGATTAAGTAAAGATACAGGT 1370
QY 1393 CTTGGACTGGCTTTCACCATTTGCCATGAGCTGACGACAACTTTGGGATGATCATGAT 1452
DB 1371 CTTGGACTGGCTTTCACCATTTGCCATGAGCTGACGACAACTTTGGGATGATCATGAT 1430
QY 1453 GGAGAAGGAAACATGTGTAAAAAGTCCGAGGCAACATCATGTCCCTACATTTGGCAGA 1512
DB 1431 GGAGAAGGAAACATGTGTAAAAAGTCCGAGGCAACATCATGTCCCTACATTTGGCAGA 1490
QY 1513 CGCAATGGAGTCTTCTCTGGTCACTTGCAGCGCCAGTATCTACACAAATTTCTAAGC 1572
DB 1491 CGCAATGGAGTCTTCTCTGGTCACTTGCAGCGCCAGTATCTACACAAATTTCTAAGC 1550
QY 1573 ACGCTCAGCTATCTGCTTCTGATCAGCCAAAGCCCTGTGAAGAAATACAGTATCTCT 1632
DB 1551 ACGCTCAGCTATCTGCTTCTGATCAGCCAAAGCCCTGTGAAGAAATACAGTATCTCT 1610
QY 1633 GAGAAATTCGCCAGGAGAAATATATGATGCAAAACACACAGTGAAGTGGCAGTTCCGAGAG 1692
DB 1611 GAGAAATTCGCCAGGAGAAATATATGATGCAAAACACACAGTGAAGTGGCAGTTCCGAGAG 1670
QY 1693 AAGCCCAAGCTCTGCATGCTGCACTTTAAAGGACATCTGTAAAGCCCTGTGGTCCCAT 1752
DB 1671 AAGCCCAAGCTCTGCATGCTGCACTTTAAAGGACATCTGTAAAGCCCTGTGGTCCCAT 1730
QY 1753 CGTATTGGAGGAAATGTGAGACTAAATTTATGCGAGCAGAGAGGACAAATTTGTGGG 1812
DB 1731 CGTATTGGAGGAAATGTGAGACTAAATTTATGCGAGCAGAGAGGACAAATTTGTGGG 1790
QY 1813 CATGACATGTGGTCCGGGAGGACAGTGTGAAATATGATGATGATGATGATGATGATGATG 1872
DB 1791 CATGACATGTGGTCCGGGAGGACAGTGTGAAATATGATGATGATGATGATGATGATGATG 1850
QY 1873 ACCATGGCCACTGTGTCGAGCTGGTCTTCTTGGTCCCATGCTCCAGGACCTGCGGAGGG 1932

1509	Db	TTGCTGATCAGCCAAAGCCTGTGAAGGAATACAAAGTATCTCTGAGAAATTTGCCAGGAGAAT	1568
1652	Qy	TATATGTCGAPACACACAGTGCAGTTCGAGGAGAGAAAGCAAAGCTTCGATGC	1711
1569	Db	TATATGTCGAAACACACAGTGCAGTTCGAGGAGAGAAAGCAAAGCTTCGATGC	1628
1712	Qy	TGGACTTTAAAAAGGACATCTGPAAGCCCTCTGGTGCCATCGTATGGAAGAAATGTG	1771
1629	Db	TGGACTTTAAAAAGGACATCTGPAAGCCCTCTGGTGCCATCGTATGGAAGAAATGTG	1688
1772	Qy	AGACTAAATTTATGCCAGCAGCAGGAGGACAAATTTCTGGCGATGACATCTGCTGCCGGS	1831
1689	Db	AGACTAAATTTATGCCAGCAGCAGGAGGACAAATTTCTGGCGATGACATCTGCTGCCGGS	1745
1832	Qy	GAGGACAGTGTGAAATATGSGTATGAAGGCCCAAGCCCAACCCATGGCCACTGGTCGS	1891
1746	Db	GAGGACAGTGTGAAATATGSGTATGAAGGCCCAAGCCCAACCCATGGCCACTGGTCGS	1805
1892	Qy	ACTGGTCTCTTGGTCCCATGTCTCGAGACCTGCGGAGGGGAGTATCTCATAGAGTCT	1951
1806	Db	ACTGGTCTCTTGGTCCCATGTCTCGAGACCTGCGGAGGGGAGTATCTCATAGAGTCT	1865
1952	Qy	GC--CTCTGCACCAACCCCAAGCCATCGCATCGAGGGAAGTCTGTGAGGGCTCCACTC	2008
1866	Db	GCCTCBAANAATACATTTCCAGGCCATCGCATGAGGGAAGTCTGTGAGGGCTCCACTC	1925
2009	Qy	GCACCTGGAAGCTCTGCAAACAGTCAGAAATGTGTCCCTGGGACAGTGTGACTTCCTGCTG	2068
1926	Db	GCACCTGGAAGCTCTGCAAACAGTCAGAAATGTGTCCCTGGGACAGTGTGACTTCCTGCTG	1985
2069	Qy	CTCAGTCTGCCGAGCACAAACAGCAGACAGATTCTCAGAGGGCGGCACTCAAGTGTGAAGCTT	2128
1986	Db	CTCAGTCTGCCGAGCACAAACAGCAGACAGATTCTCAGAGGGCGGCACTCAAGTGTGAAGCTT	2044
2129	Qy	ACACTCAAGTAGAAGATCAGACATTATGCAAACTCTACTGTATGCGAAGAAGTATTTGATT	2188
2045	Db	-----GATCAGGACTTATGCAAACTCTACTGTATGCGAAGAAGTATTTGATT	2090
2189	Qy	TCCTCTTTCTTTCTGCAAAATGAAGTGAAGTGGACTCCATGCTCGGAGGATAGCGTA	2248
2091	Db	TCCTCTTTCTTTCTGCAAAATGAAGTGAAGTGGACTCCATGCTCGGAGGATAGCGTA	2150
2249	Qy	ATGTTCTATAGATGGATATGTGAGAGAGTTGGATGTGCAATGTCTTGGATCTGATG	2308
2151	Db	ATGTTCTATAGATGGATATGTGAGAGAGTTGGATGTGCAATGTCTTGGATCTGATG	2210
2309	Qy	CTGTTGAGAGCTCTGTGGGTGTGAACGGGAATACTCAGGCTGCACCATTCACAGGG	2368
2211	Db	CTGTTGAGAGCTCTGTGGGTGTGAACGGGAATACTCAGGCTGCACCATTCACAGGG	2270
2369	Qy	GTCTCTACCAAGCACCAACACACCAAGTATTTATCATGGTGCACCATCTCTCTG	2428
2271	Db	GTCTCTA-----CCTAGAGTATTTATCATGGTGCACCATCTCTCTG	2312
2429	Qy	GAGCCCGAGTATCCGATCTATGAATGACGCTCTACTCTCTACATTTCTGTGGCA	2488
2313	Db	GAGCCCGAGTATCCGATCTATGAATGACGCTCTACTCTCTACATTTCTGTGGCA	2372
2489	Qy	ATGCCCTCAGAAGTACTACTCTGAATGGGCACTGGACGTGGACTGCCCGGCGGTACA	2548
2373	Db	ATGCCCTCAGAAGTACTACTCTGAATGGGCACTGGACGTGGACTGCCCGGCGGTACA	2432
2549	Qy	AATTTTGGGGCACTACTTTTGACATACAGACGGTCTCTAATAGGCCCGAGAACTTAATCG	2608
2433	Db	AATTTTGGGGCACTACTTTTGACATACAGACGGTCTCTAATAGGCCCGAGAACTTAATCG	2492
2609	Qy	CTACTGGACCAACCAAGAGACACTGATTGTGGAGCTGCTGTTTTCAGGGAAGGAACCCGG	2668
2493	Db	CTACTGGACCAACCAAGAGACACTGATTGTGGAGCTGCTGTTTTCAGGGAAGGAACCCGG	2552
2669	Qy	GTGTTGCTGGGAATATCTCAATGCTCTGCTTTGGGACCGAGAACGCCCTTCCGACG	2728

Db	2553	GTGTTGCCCTGGGAATCACTCCATGCTCTGGCTGGGGAGCAGAGACGCCCTTGCCTGCACG	2611
Qy	2729	CCAGCTACACTTGGGCGCATCGTGGCTCTGAGTGTCTCCGTGCTCCGTGCGAGGGGTAGGT	2788
Db	2613	CCAGCTACACTTGGGCGCATCGTGGCTCTGAGTGTCTCCGTGCTCCGTGCGAGGGGTAGGT	2672
Qy	2789	GCCTTCAGTGTCTCTGGAGGAGCATGTCTCAGCCTTTCAGCCATGCGTACATGTCAC	2848
Db	2673	GCCTTCAGTGTCTCTGGAGGAGCATGTCTCAGCCTTTCAGCCATGCGTACATGTCAC	2732
Qy	2849	TGCCTCTTCTTGAATCCTAATGAGCAGCCGGGGCTTCTCCCTGCCA	2895
Db	2733	TGCCTCTTCTTGAATCCTAATGAGCAGCCGGGGCTTCTCCCTGCCA	2779
RESULT 13			
LOCUS	BD170079	Novel ADAMTS family polypeptide and gene encoding the same.	
DEFINITION	BD170079	Novel ADAMTS family polypeptide and gene encoding the same.	
ACCESSION	BD170079	Novel ADAMTS family polypeptide and gene encoding the same.	
VERSION	BD170079.1	GI:27875891	
KEYWORDS	WO 0231163-A/1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 4234) Chara, O., Nagase, T., Nomura, N., Yano, K., Murakami, K., Yasuda, S. and Kanzaki, K.		
TITLE	Novel ADAMTS family polypeptide and gene encoding the same		
JOURNAL	Patent: WO 0231163-A 1 18-APR-2002; KAZUSA DNA RESEARCH INSTITUTE, MITSUBISHI PHARMA CORP, OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO YASUDA, KOJI KANZAKI		
COMMENT	OS Homo sapiens (human) PN WO 0231163-A/1 PD 18-APR-2002 PF 11-OCT-2001 WO 2001JP008913 PR 11-OCT-2000 JP 00P 311309, 02-APR-2001 JP 01P 102905 PI OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO YASUDA, KOJI KANZAKI PI SHINICHIRO YASUDA, KOJI KANZAKI PC C12N15/57, C12N9/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/16 PC 40, C12Q1/68, PC G01N33/50 CC Novel ADAMTS family polypeptide and gene encoding the same FH Key Location/Qualifiers FT CDS Location/Qualifiers I. 4234 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
BASE COUNT	1036 a 1103 c 1130 g 965 t		
ORIGIN			
Query Match	56.5%	Score 1634.8; DB 6; Length 4234;	
Best Local Similarity	93.3%	Prod. No. 0;	
Matches 1772;	Conservative	0; Mismatches 2; Indels 126; Gaps 1;	
Qy	1010	ACATGCCCCAGCCTCCCAAGGAAGACCTCTTCATCTTGCCAGATGAGTAAAGCTTGTGCT	1069
Db	157	ACATGCCCCAGCCTCCCAAGGAAGACCTCTTCATCTTGCCAGATGAGTAAAGCTTGTGCT	216
Qy	1070	TACGGCATAAAGCGCTCTCTTCTGAGTCCCATAGAAATGAAGAACTGAACGTCGACACCT	1129
Db	217	TACGGCATAAAGCGCTCTCTTCTGAGGTACCATAGAAATGAAGAACTGAACGTCGACACCT	276
Qy	1130	TGTTGGTGTGCGACAAAGATGATCGAACCATTGCGCCATGAAATATCACCACCTAGC	1189
Db	277	TGTTGGTGTGCGACAAAGATGATCGAACCATTGCGCCATGAAATATCACCACCTAGC	336
Qy	1190	TGTTCCAGCATCTCAACATGGTATCTGCTTTATTCAAGA-	1229

Db 337 TGCACGATCTCAACATGCTGCTTTTATTCAGAGATGAACATAGGAGAAACA 396
QY 1230 ----- 1229
Db 397 TCACATTTGCAATTTGATGCTGATTTCTTAGAAGATGAACAGCCAGGACTGGTGATAA 456
QY 1230 -----TGATTTGATGGGGA 1243
Db 457 GTCACACGCGACACACACTTAAGTAGTTCTGCCAGTGGCAGTCTGGATGATGGGGA 516
QY 1244 AAGATGGGACTCGTCATGACACCGCCATCTTACTGACTGGTCTGGATATATGTTCTCTGGA 1303
Db 517 AAGATGGGACTCGTCATGACACCGCACTTACTGACTGGTCTGGATATATGTTCTCTGGA 576
QY 1304 AAGATGGGACTCGTCATGACACCGCACTTACTGACTGGTCTGGATATATGTTCTCTGGA 1363
Db 577 AAGATGGGACTCGTCATGACACCGCACTTACTGACTGGTCTGGATATATGTTCTCTGGA 636
QY 1364 GCAGCTGCACGATTAATGAAGATACAGGCTTTGGACTGGCCCTTCACCAATTCGCCATGAGT 1423
Db 637 GCAGCTGCACGATTAATGAAGATACAGGCTTTGGACTGGCCCTTCACCAATTCGCCATGAGT 696
QY 1424 CTGGACACAACTTTGGCAGTATTCATGATGGAGAGGGAACATGTTTAAAAAGTCCGAGG 1483
Db 697 CTGGACACAACTTTGGCAGTATTCATGATGGAGAGGGAACATGTTTAAAAAGTCCGAGG 756
QY 1484 GCAACATATGTCCTTACATTTGGCAGGAGCAATGAGTCTTCTCTGCTGGTCAACCTGCA 1543
Db 757 GCAACATATGTCCTTACATTTGGCAGGAGCAATGAGTCTTCTCTGCTGGTCAACCTGCA 816
QY 1544 GCCGCGAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTGCTGATCAGC 1603
Db 817 GCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTGCTGATCAGC 876
QY 1604 CAAAGCTGTGAAGGAATCAAGTATCTTGAAGAAATGGCAGAGAAATATATGATGCA 1663
Db 877 CAAAGCTGTGAAGGAATCAAGTATCTTGAAGAAATGGCAGAGAAATATATGATGCA 936
QY 1664 ACACAGTGCAGTGGGAGTCTGGAGAGAAAGCCAAAGCTCTGATGCTGGACTTTAAAA 1723
Db 937 ACACAGTGCAGTGGGAGTCTGGAGAGAAAGCCAAAGCTCTGATGCTGGACTTTAAAA 996
QY 1724 AGGACATCTGTAAGCCCTGTGGTGCATCGTATTGGAAGGAATATGAGACTAAATTTA 1783
Db 997 AGGACATCTGTAAGCCCTGTGGTGCATCGTATTGGAAGGAATATGAGACTAAATTTA 1056
QY 1784 TGCCAGCAGCAGAGGACAAATTTGTGGGATGACATGTGTGCGCCGAGGAGGACAGTGTG 1843
Db 1057 TGCCAGCAGCAGAGGACAAATTTGTGGGATGACATGTGTGCGCCGAGGAGGACAGTGTG 1116
QY 1844 TGAATATGTGATGAAGGCCCAAGCCCAAGCCATGCGCACTGGTGGGACTGGTCTTCT 1903
Db 1117 TGAATATGTGATGAAGGCCCAAGCCCAAGCCATGCGCACTGGTGGGACTGGTCTTCT 1176
QY 1904 GGTCCCATGCTCCAGGACTCGGAGGGGAGTATCTCATAGAGTGCCTCTGCACCA 1963
Db 1177 GGTCCCATGCTCCAGGACTCGGAGGGGAGTATCTCATAGAGTGCCTCTGCACCA 1236
QY 1964 ACCCAAGCCATGCAATGGAGGAAGTCTGTGAGGGCTCCACTCCGACTCTGAAGCTCT 2023
Db 1237 ACCCAAGCCATGCAATGGAGGAAGTCTGTGAGGGCTCCACTCCGACTCTGAAGCTCT 1296
QY 2024 GCAACAGTCAAGAAATGCCCGGAGCAGTGTGACATTCCTGCTGCTGCTGCTGCTGCTG 2083
Db 1297 GCAACAGTCAAGAAATGCCCGGAGCAGTGTGACATTCCTGCTGCTGCTGCTGCTGCTG 1356
QY 2084 ACAACAGCAGACGATTCAGAGGGCGGCACTTCAAGTGGAGGCTTACACTCAAGTAGAAG 2143
Db 1357 ACAACAGCAGACGATTCAGAGGGCGGCACTTCAAGTGGAGGCTTACACTCAAGTAGAAG 1416
QY 2144 ATCAGGACTTATGCAAACTCTACTGATTCGACAGAGATTTGATTTCTTTCTTTCTTGT 2203

Db 1417 ATCAGGACTTATGCAAACTCTACTGTATCGAGAGGATTTGATTTCTTCTTTCTTGT 1476
QY 2204 CAAATAAGTCAAAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATG 2263
Db 1477 CAAATAAGTCAAAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATG 1536
QY 2264 GGATATGTGAGAGAGTTGGATGTGACAAATGTCTTGGATCTGATGCTGTTGGAAGACGTCT 2323
Db 1537 GGATATGTGAGAGAGTTGGATGTGACAAATGTCTTGGATCTGATGCTGTTGGAAGACGTCT 1596
QY 2324 GTGGGCTGTAAACGGGAATTAACCTCAGCTGCAGATTCACAGGGGTCTCTACACCAAGC 2383
Db 1597 GTGGGCTGTAAACGGGAATTAACCTCAGCTGCAGATTCACAGGGGTCTCTACACCAAGC 1656
QY 2384 ACCACCAACCAACAGATTTATCAGATGCTGACATTTCTTGGAGCCCGGAGTATCC 2443
Db 1657 ACCACCAACCAACAGATTTATCAGATGCTGACATTTCTTGGAGCCCGGAGTATCC 1716
QY 2444 GCATCTATGAAATGAAACGCTCTTACCTCTTACCTCTTCTGCGCAATGCCCTCAGAGGT 2503
Db 1717 GCATCTATGAAATGAAACGCTCTTACCTCTTACCTCTTCTGCGCAATGCCCTCAGAGGT 1776
QY 2504 ACTACCTGAATGGGACTTGGACCTGAGCTGGCCCGGCGGTACAAATTTTTCGGGCACTA 2563
Db 1777 ACTACCTGAATGGGACTTGGACCTGAGCTGGCCCGGCGGTACAAATTTTTCGGGCACTA 1836
QY 2564 CTTTCGACTACAGACGGTCTTATATGAGCCCGGAGAACTTAACTCGTACTAGGACCAACA 2623
Db 1837 CTTTCGACTACAGACGGTCTTATATGAGCCCGGAGAACTTAACTCGTACTAGGACCAACA 1896
QY 2624 ACGAGACACTGATTTGTGAGCTGCTTTCAGGGAGGAAACCCGGGTGTTCCTGGGAAAT 2683
Db 1897 ACGAGACACTGATTTGTGAGCTGCTTTCAGGGAGGAAACCCGGGTGTTCCTGGGAAAT 1956
QY 2684 ACTCCATGCTGCTTGGGAGCCGAGAGAGCCCTGCGCCAGCCAGCTACACTTGGG 2743
Db 1957 ACTCCATGCTGCTTGGGAGCCGAGAGAGCCCTGCGCCAGCCAGCTACACTTGGG 2016
QY 2744 CCATCGTGGCTCTGAGTGTCTCGTGTCTCGGAGGGGG 2783
Db 2017 CCATCGTGGCTCTGAGTGTCTCGTGTCTCGGAGGGGG 2056

RESULT 14

AB095949
LOCUS AB095949 4234 bp mRNA linear PRI 12-NOV-2002
DEFINITION Homo sapiens mRNA for KIAA2029 protein.

ACCESSION

AB095949

VERSION

AB095949.1 GI:24899221

KEYWORDS

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 Nagase, T., Kikuno, R. and Ohara, O.

The nucleotide sequence of a long cDNA clone isolated from human

Published Only in Database (2002)

JOURNAL

2 (bases 1 to 4234)

REFERENCE

Nagase, T., Kikuno, R. and Ohara, O.

AUTHORS

Direct Submission

TITLE

Submitted (08-NOV-2002) Osamu Ohara, Kazusa DNA Research Institute,

Department of Human Gene Research; 2-6-7 Kazusa-kamata, Kisarazu,

Chiba 292-0818, Japan (E-mail: cdnainfo@kazusa.or.jp,

URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,

Fax: 81-438-52-3914)

Location/Qualifiers

1. .4234

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="pj01256"

/tissue_type="Brain"

/note="vector:pBluescript II SK plus"		
gene	1. 4234	
CDS	/gene="KIAA2029"	
	<1. 3068	
	/gene="KIAA2029"	
	/codon_start=3	
	/evidence=not_experimental	
	/product="KIAA2029 protein"	
	/protein_id="BAC23125.1"	
	/db_xref="GI:2489922"	
	/translation="LYKSTEPHAPAGASEVLVTSRTWELAHQPHLSDLRLGLPQKH FCGRKXKMPQPKEDFILPDEVKSLRKRLRLRHEELNVEVLVVDKMMQON HGHENITVLTILNMYSAIPKDGITGNINLAIWGLLLLEDEQGLVI SHHADHTLS SFCQWQSLMGKDGTRHDHAILLGLDICSWNPECDTLGFAPISGSKVRSCTINE DTGLGFTAI AHESGHNFMIHDGEMCKSEGNIMSPITLADNGNNGVFSWPSRQYL HKFLSTAQAI CLADQPKPVKEYKPEKLPGLYDANTCKWQFGEKAKLMLDFPKDI CKALWCHRI HKGKCTKMPAAEAGTICGHDMCRGGQVKYGDGEPKTHGHSWSSN SPCSRTCGGVSHKSLCTNPKPSHGKFCBGSRTRLKLNQSKRDKDTPCSEDSRVV EHNRRPRGRHKVPYTOVEDQDLCKLYCIAEGFDFFFLSNKVKDGTTCSEDSRVV CIDGICERVGCDNLGSDAEDVGVGCGNNSACTIHRGLYTKHHHTNOYVHMVTPS GARSIRIYEMNVSTSYISVRNALRYILNHWTDWPGRYKESCTTIDYERSVNEPEN LIATGPTNETLIVLELLFGQNPWAWSEYMPRLGTEKOPPAQPSYITWAI VRSECSVC GGQMTVREGYRLDLKFQVNNFCNPKTPVTGLVPCVKSACPSWSVGNWMSCSRTC GGAGQRPVQCTRRVHDSPEVPSALCPQAPSPSRQACNSQSCPPASGAPWAECSRT CGKGRKBAVACKSNPSARQLLPDAVCTSEPKRMHEACLLQKCHKPKKLQWLVA WSQSVTCERGTRKPLCAEKYVSGKVKRELASKSHLKPESLEBRACAPLPCPRH PPFAAGAPSRGSPFASPMWCTASCGGVQVTRSVQCLAGRPASGCLLHQKPSASLAC NTHFCPIAEKDAFCDFYHMCYDLVPOHGMCHKPYGKQCCCKSKSL"	
BASE COUNT	1036 a 1103 c 1130 g 965 t	
ORIGIN	56.5%; Score 1634.8; DB 9; Length 4234; Query Match 93.3%; Pred. No. 0; Mismatches 126; Gaps 1; Matches 1772; Conservative 0;	
QY	1010	ACATGCCCGAGCTCCCAAGGAAGACCTCTTCATCTTGCAGATGAGTATAAGTCTTCT 1069
DB	157	ACATGCCCGAGCTCCCAAGGAAGACCTCTTCATCTTGCAGATGAGTATAAGTCTTCT 216
QY	1070	TACGGGATTAAGGCTCTCTTCTGAGTCCCATGAGAAATGAGAACTGACAGTGGAGACCT 1129
DB	217	TACGGGATTAAGGCTCTCTTCTGAGTACCATAGAAATGAGAACTGACAGTGGAGACCT 276
QY	1130	TGGTGGTGCACAAAAGAGATGATGCAAAACCATGGCCATGAAATATCAACCACTACG 1189
DB	277	TGGTGGTGCACAAAAGAGATGATGCAAAACCATGGCCATGAAATATCAACCACTACG 336
QY	1190	TGCTACGATACTCAACATGGTATCTGCTTTATTCAAAGA----- 1229
DB	337	TGCTACGATACTCAACATGGTATCTGCTTTATTCAAAGA----- 1229
QY	1230	----- 1229
DB	397	TCAACATATGCAATGTAGTCTGATTCCTTAGAAGATGACAGCAGCAGCTGCTGATAA 456
QY	1230	-----TGATTTGATGGGA 1243
DB	457	GTCAACACGAGACACACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGA 516
QY	1244	AGATGGGACTCGTCTATGACACGCCATCTTACTGACTGCTGGATATATCTTCTCTGA 1303
DB	517	AAGATGGGACTCGTCTATGACACGCCATCTTACTGACTGCTGGATATATGTTCTCTGA 576
QY	1304	AGATGGGCTCTGACACTTTTGGATTTCACCCATAGTGAATAGTGTAGTAAATATC 1363
DB	577	AGATGGGCTCTGACACTTTTGGATTTCACCCATAGTGAATAGTGTAGTAAATATC 636
QY	1364	GCAGCTGACGATTAAATGAGATACAGGCTCTGAGCTGGCCCTTCAACATGGCCATGAGT 1423
DB	637	GCAGCTGACGATTAAATGAGATACAGGCTCTTGGACTGGCCCTTCAACATGGCCATGAGT 696
QY	1424	CTGGACACAACTTTGGATGATTCATGATGAGAGGACATGTGTAAAGTCCGAGG 1483

697	CTGGACACAACTTTGGCATGATTCATGATGGAGAGGGAACATGTGCAAAAAATCCGAGG	756
1484	GCACATCATGTCCTCCCTACATATTGGCAGGACGCAATCGAGTCTTCTCCTGGTCACTCCGCA	1543
757	GCAACATCATGTCCTCCCTACATATTGGCAGGACGCAATCGAGTCTTCTCCTGGTCACTCCGCA	816
1544	GCCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCCTGTGTGATCAGC	1603
817	GCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCCTGTGTGATCAGC	876
1604	CAAAAGCTGTGAAGGAATACAGATATCTTGAGAAATTCAGGAGAAATATATATGATGCA	1663
877	CAAAAGCTGTGAAGGAATACAGATATCTTGAGAAATTCAGGAGAAATATATATGATGCA	936
1664	ACACAGTGTCAAGTGGCAGTTCGGAGAGAAAGCCAGCTCTCATGCTGGAGCTTTAAAA	1723
937	ACACAGTGTCAAGTGGCAGTTCGGAGAGAAAGCCAGCTCTCATGCTGGAGCTTTAAAA	996
1724	AGGACATCTCTAAAGCCCTGTGTGTCATCTGTATTTGAAGAAATGTGAGACTAAATTTA	1783
997	AGGACATCTCTAAAGCCCTGTGTGTCATCTGTATTTGAAGAAATGTGAGACTAAATTTA	1056
1784	TGCAGCAGCAGAGGACACAAATTTGTGGGCATGACATGTGTGCGGGAGGACAGTGTG	1843
1057	TGCAGCAGCAGAGGACACAAATTTGTGGGCATGACATGTGTGCGGGAGGACAGTGTG	1116
1844	TGAAATATGTGATGAAGGCCCAAGCCCAATGCGCACTGTGTGCGACTGTGTCTTCTT	1903
1117	TGAAATATGTGATGAAGGCCCAAGCCCAATGCGCACTGTGTGCGACTGTGTCTTCTT	1176
1904	GGTCCCCTATGCTCAGACCTTGCAGGGGGAGTATCTCATAGAGTGGCTCTGCACCA	1963
1177	GGTCCCCTATGCTCAGACCTTGCAGGGGGAGTATCTCATAGAGTGGCTCTGCACCA	1236
1964	ACCCCAAGCCATCCCATGAGGGGAGTTCGTGAGAGGCTCCACCTCGCACTCTGAAAGTCT	2023
1237	ACCCCAAGCCATCCCATGAGGGGAGTTCGTGAGAGGCTCCACCTCGCACTCTGAAAGTCT	1296
2024	GCAACAGTCAGAAATGTCCCCGGGACAGTGTGTACTTCCGCTGCTGCTGAGTGGCAGC	2083
1297	GCAACAGTCAGAAATGTCCCCGGGACAGTGTGTACTTCCGCTGCTGCTGAGTGGCAGC	1356
2084	ACAACAGCAGAGATTCAGAGGGGGGACACTACAGTGGAGCCCTTACACTCAAGTAGAAG	2143
1357	ACAACAGCAGAGATTCAGAGGGGGGACACTACAGTGGAGCCCTTACACTCAAGTAGAAG	1416
2144	ATCAGGACTTATGCAAACTCTACTGTATCGCAGAGAGATTTGATTTCTTTCTTTTGT	2203
1417	ATCAGGACTTATGCAAACTCTACTGTATCGCAGAGAGATTTGATTTCTTTCTTTTGT	1476
2204	CAAAATAAGTCAAAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATG	2263
1477	CAAAATAAGTCAAAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATG	1536
2264	GGATATGTGACAGAGTTCGATGTCACATGCTCTTGGATCTGATCTGTGAAGACGCT	2323
1537	GGATATGTGACAGAGTTCGATGTCACATGCTCTTGGATCTGATCTGTGAAGACGCT	1596
2324	GTGGGGTGTAAACCGGGAATAACTCAGCTTCGACGATTCACAGGGGCTCTTACACCAAGC	2383
1597	GTGGGGTGTGTAAACCGGGAATAACTCAGCTTCGACGATTCACAGGGGCTCTTACACCAAGC	1656
2384	ACCACCAACCAACAGATTTATCATGTTGTCATCATCTTCTTGGAGCCCGGATATCC	2443
1657	ACCACCAACCAACAGATTTATCATGTTGTCATCATCTTCTTGGAGCCCGGATATCC	1716
2444	GCATCTATGAAATGAACGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2503
1717	GCATCTATGAAATGAACGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1776
2504	ACTTACTGAAATGGGACCTGGACCGTGGAGTGGCCCGCCGCGGTACAAATTTTCGGSCACTA	2563
1777	ACTTACTGAAATGGGACCTGGACCGTGGAGTGGCCCGCCGCGGTACAAATTTTCGGSCACTA	1836


```
|||||
Db 5306 AGATCAGGACTTATGCAAACTCTACTGTAATCGAAGGATTTGATTTCTCTTTCTTT 5365
QY 2202 GTCAATAAAGTCAAAGATGGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATAGA 2261
Db 5366 GTCAATAAAGTCAAAGATGGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATAGA 5425
QY 2262 TGGGATATGTGAGAGGTTGGATGTGACAAATGTCCTTGGATCTGATGCTGTTGAAGAGT 2321
Db 5426 TGGGATATGTGAGAGGTTGGATGTGACAAATGTCCTTGGATCTGATGCTGTTGAAGAGT 5485
QY 2322 CTGTGGGGTGTAAACGGGAATAACTCAGCTGCAGATTACAGGGGTCTCTACACCAA 2381
Db 5486 CTGTGGGGTGTAAACGGGAATAACTCAGCTGCAGATTACAGGGGTCTCTACACCAA 5545
QY 2382 GCACCAACCAACCAACAGTATTATCAGATGTCACCAATTCCTTCTGGAGCCCGAGTAT 2441
Db 5546 GCACCAACCAACCAACAGTATTATCAGATGTCACCAATTCCTTCTGGAGCCCGAGTAT 5605
QY 2442 CCGCATCTATGAATGAAGTCTCTACTCTCTACATTTCTGTGCGCAATGCCCTCAGAG 2501
Db 5606 CCGCATCTATGAATGAAGTCTCTACTCTCTACATTTCTGTGCGCAATGCCCTCAGAG 5665
QY 2502 GTACTACCTGAATGGGCACCTGGACCGTGGACTGGCCCGCGGTACAAATTTTCGGGCAC 2561
Db 5666 GTACTACCTGAATGGGCACCTGGACCGTGGACTGGCCCGCGGTACAAATTTTCGGGCAC 5725
QY 2562 TACTTTCGACTACAGACGGTCTTAATAGAGCCCGAGAACTTAATCGCTACTGGACCAAC 2621
Db 5726 TACTTTCGACTACAGACGGTCTTAATAGAGCCCGAGAACTTAATCGCTACTGGACCAAC 5785
QY 2622 CAACGAGACACTGATTGTGGAGTGTCTGTTTCAGGGAAGAAACCCGGGTGTTGCCGTGGGA 2681
Db 5786 CAACGAGACACTGATTGTGGAGTGTCTGTTTCAGGGAAGAAACCCGGGTGTTGCCGTGGGA 5845
QY 2682 ATACTCCATGCCCTCGCTTGGGGAACCGAGAGAGCCCTTGCCTCCAGCCCGAGCTACACTTG 2741
Db 5846 ATACTCCATGCCCTCGCTTGGGGAACCGAGAGAGCCCTTGCCTCCAGCCCGAGCTACACTTG 5905
QY 2742 GGCCTATCGTGGCTCTGAGTGTCTCCGTGCTCCTCGGAGGGG 2783
Db 5906 GGCCTATCGTGGCTCTGAGTGTCTCCGTGCTCCTCGGAGGGG 5947
```

Search completed: October 28, 2003, 21:53:52
Job time : 10283 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2003, 17:18:08 ; Search time 726 Seconds
(without alignments)
10764.288 Million cell updates/sec

Title: US-09-981-151A-7

Perfect score: 2895

Sequence: 1 cgcctcgtgatgaagcccg.....ccgggggtcttcctcgcca 2895

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2895	100.0	2895	24	ABS59326
2	2675.4	92.4	2997	24	ABS59323
3	1877.8	64.9	3675	24	ABS97174
4	1877.8	64.9	5610	24	ABS97176
5	1877.8	64.9	8435	24	ABK49822
6	1877.8	64.9	8505	24	ABK49825
7	1874.6	64.8	3675	24	AAU43654
8	1874.6	64.8	3675	25	ABS57767

9	1874.6	64.8	4042	25	ABS57768	cDNA encoding nove
10	1783.6	61.6	2432	24	ABS59324	Human zinc matallo
11	1695	58.5	2902	24	ABS59325	Human ADAM-TS 7-li
12	1638	56.6	4888	24	ABK86137	cDNA encoding huma
13	1634.8	56.5	4234	24	ABK49821	Human cDNA encodin
14	1578.4	54.5	7668	24	ABK49827	Plasmid pFastBac1-
15	1381.8	47.7	1668	24	ABK90322	DNA encoding human
16	708	24.5	3666	24	ABN92777	Human aggrecanase
17	706.4	24.4	2805	24	AS97176	Human metalloprote
18	678.6	23.4	3830	25	ABT16078	NOVX related polyn
19	673.4	23.3	2589	25	ABZ79548	Human protease (NH
20	673.4	23.3	3013	25	ABZ79549	Human protease (NH
21	663	22.9	1476	25	ABS57766	cDNA encoding nove
22	434.4	15.0	1059	25	ABT16081	NOVX related polyn
23	434.4	15.0	1059	25	ABT16083	NOVX related polyn
24	432.8	14.9	1059	25	ABT16079	NOVX related polyn
25	432.8	14.9	1059	25	ABT16080	NOVX related polyn
26	432.8	14.9	1059	25	ABT16082	NOVX related polyn
27	417.6	14.4	545	24	ABK49824	Mouse cDNA encodin
28	356.6	12.3	785	25	ACC46465	Human dithp protei
29	320.2	11.1	422	24	ABK49823	Mouse cDNA encodin
30	241.2	8.3	2217	25	ABZ79551	Human protease (NH
31	227	7.8	2274	22	AAO07967	Human protease-rel
32	227	7.8	2727	22	AAO07956	Human protease-rel
33	227	7.8	2879	22	AAH49370	Human metalloprote
34	225.4	7.4	5338	25	ABZ79550	Human metalloprote
35	215.4	7.4	5338	22	ABH49372	Human metalloprote
36	212.2	7.3	5353	22	AAO04494	Human metalloprote
37	211.8	7.3	5353	22	AAO04475	Human 27875 ADAM-T
38	211.2	7.3	6565	24	ABS58373	Protein modificati
39	210.6	7.3	5061	22	AAF82149	Human ADAM type me
40	209.6	7.2	3571	24	ABZ11529	Human polynucleoti
41	208	7.2	3133	24	ABA94919	Human metalloprote
42	208	7.2	3312	22	AAF82157	Human ADAM type me
43	208	7.2	3312	24	AS97173	Human metalloprote
44	208	7.2	3329	24	ABQ78455	Nucleotide sequenc
45	208	7.2	3329	24	ABA94918	Human metalloprote

ALIGNMENTS

RESULT 1

ABS59326

ID ABS59326 standard; DNA; 2895 BP.

XX AC ABS59326;

XX AC ABS59326;

DT 05-NOV-2002 (first entry)

XX DE Human ADAM-TS 7-like gene #2.

XX DE Human ADAM-TS 7-like gene #2.

KW Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing;
KW breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
KW anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
KW neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
KW autoimmune disease; allergy; addiction; asthma; transplantation;
KW graft versus host disease; systemic lupus erythematosus; scleroderma;
KW psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
KW atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
KW thrombocytopenia; bleeding disorder; glomerulonephritis; hypercalcaemia;
KW polycystic kidney disease; renal tubular acidosis; skin disorder;
KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
KW haematopoietic disorder; urinary system disorder; osteoporosis; ds;
KW dental disease; dental infection; growth disorder; reproductive disorder;
KW hypogonadism; fertility disorder; viral infection; bacterial infection;
KW parasitic infection; metabolic pathway modulation; gene therapy; gene;
KW zinc metalloprotease; ADAM-TS 7; alpha-2-macroglobulin precursor;
KW ileal sodium/bile acid cotransporter; prohibitin; MT; ClpA; spinesin;
KW macrophage stimulating protein precursor; fatty acid-binding protein;
KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.

1081	CGCTCTCTTCTGAGTCCCATAGAAATGAAGAACTGAACTGGAGACCTTGGTGTGGTCT	Qy	1140
1081	CGCTCTCTTCTGAGTCCCATAGAAATGAAGAACTGAACTGGAGACCTTGGTGTGGTCT	Db	1140
1141	GACAAAAGATGATGCARAAACCATGGCCATGAAATATACACCACTACGTGCTCACGATA	Qy	1200
1141	GACAAAAGATGATGCARAAACCATGGCCATGAAATATACACCACTACGTGCTCACGATA	Db	1200
1201	CTCAACATGGTATCTGCTTTATTCAAAGATGGATTCATGGGAAAGATGGGACTCGTCAT	Qy	1260
1201	CTCAACATGGTATCTGCTTTATTCAAAGATGGATTCATGGGAAAGATGGGACTCGTCAT	Db	1260
1261	GACCAGCCCATCTTACTGACTGGTCTGGATATATGTTCTCGAAGAAATGAGCCCTGTGAC	Qy	1320
1261	GACCAGCCCATCTTACTGACTGGTCTGGATATATGTTCTCGAAGAAATGAGCCCTGTGAC	Db	1320
1321	ACTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAAT	Qy	1380
1321	ACTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAAT	Db	1380
1381	GAGATACAGGCTTTGGACTGGCTTACCATGCCCCATGAGTCTGGACACAACTTTGGC	Qy	1440
1381	GAGATACAGGCTTTGGACTGGCTTACCATGCCCCATGAGTCTGGACACAACTTTGGC	Db	1440
1441	ATGATTTCATGATGGAGAAGGAAATGTGTAAAGTCCGAGGGCAACATCATGTCCCT	Qy	1500
1441	ATGATTTCATGATGGAGAAGGAAATGTGTAAAGTCCGAGGGCAACATCATGTCCCT	Db	1500
1501	ACATTCGAGGACGGCAATGGAGTCTCTCTGTGTACCTCGAGCCCGCAGTATCTACAC	Qy	1560
1501	ACATTCGAGGACGGCAATGGAGTCTCTCTGTGTACCTCGAGCCCGCAGTATCTACAC	Db	1560
1561	AAATTTCTAAGCACCGCTCAAGCTATCTGCCTGTCTGATCAGCCAAAGCTGTGAAGGAA	Qy	1620
1561	AAATTTCTAAGCACCGCTCAAGCTATCTGCCTGTCTGATCAGCCAAAGCTGTGAAGGAA	Db	1620
1621	TACAGTATCTCTGAGAAATGCCAGAGAAATATATGNTGCMAACACACAGTSCAAAGTG	Qy	1680
1621	TACAGTATCTCTGAGAAATGCCAGAGAAATATATGNTGCMAACACACAGTSCAAAGTG	Db	1680
1681	CAGTTCGGAGAGAAAGCCAGCTCTGCATCTGGACTTTTAAAGAGACATCTGTAAAGCC	Qy	1740
1681	CAGTTCGGAGAGAAAGCCAGCTCTGCATCTGGACTTTTAAAGAGACATCTGTAAAGCC	Db	1740
1741	CTGTGTGCCATCGTATTTGGAAGAAATGTGAGACTAAATTTATTCGACAGCAGAGAGC	Qy	1800
1741	CTGTGTGCCATCGTATTTGGAAGAAATGTGAGACTAAATTTATTCGACAGCAGAGAGC	Db	1800
1801	ACAATTTGGGCATGCACATGTGGTCCGGGGAGACAGTGTGTGAAATATGGTGATGAA	Qy	1860
1801	ACAATTTGGGCATGCACATGTGGTCCGGGGAGACAGTGTGTGAAATATGGTGATGAA	Db	1860
1861	GGCCCCAAGCCCATCGCCACTGGTCCGACTGGTCTTTTGTGTGCCCATGCTCCAGS	Qy	1920
1861	GGCCCCAAGCCCATCGCCACTGGTCCGACTGGTCTTTTGTGTGCCCATGCTCCAGS	Db	1920
1921	ACCTGCGAGGGGGAGTATCTCATAGAGTGGCTCTGCACCAAGCCCGAGCCATCGCAT	Qy	1980
1921	ACCTGCGAGGGGGAGTATCTCATAGAGTGGCTCTGCACCAAGCCCGAGCCATCGCAT	Db	1980
1981	GGAGGGAAGTTCTGTAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAATGT	Qy	2040
1981	GGAGGGAAGTTCTGTAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAATGT	Db	2040
2041	CCCCGGGACAGTGTGACTTTCGGTGTCTCAGTGTGCCGAGCACAAACAGCAGCAATTC	Qy	2100
2041	CCCCGGGACAGTGTGACTTTCGGTGTCTCAGTGTGCCGAGCACAAACAGCAGCAATTC	Db	2100
2101	AGAGGGCGGCATCTACAGTGGAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCBA	Qy	2160
2101	AGAGGGCGGCATCTACAGTGGAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCBA	Db	2160
2161	CTCTACTGTATCCGAGAAGGATTTGATTTCTTTCTTTTGTCAAATAAAGTCAAGAT	Qy	2220

	Db	2161	CTCTACTGTATCGCAGAGGATTGAATTTCTTCTTTCTTTGGTCAATTAAGTCAAAGAT	2220
	Qy	2221	GGGACTCCCATGCTCGGAGGATAGCCGTAATGTTTGATAGATGGGATATGTGAGAGAGTT	2280
	Db	2221	GGGACTCCCATGCTCGGAGGATAGCCGTAATGTTTGATAGATGGGATATGTGAGAGAGTT	2280
	Qy	2281	GGATGTGACAAATGTCCTTGATCTGATGCTGTGTGAAGAGCGTCTGTGGGGTGTGTACCGG	2340
	Db	2281	GGATGTGACAAATGTCCTTGATCTGATGCTGTGTGAAGAGCGTCTGTGGGGTGTGTACCGG	2340
	Qy	2341	AATAAATCGAGCTGCAGATTTCACAGGGGTCTCTACACAAGCACACCACACCAACCAG	2400
	Db	2341	AATAAATCGAGCTGCAGATTTCACAGGGGTCTCTACACAAGCACACCACACCAACCAG	2400
	Qy	2401	TATTATCACATGTGTACCATTTCCTTCTGGAGCCCGGAGTATCCGCCATCTATGAAATGAAC	2460
	Db	2401	TATTATCACATGTGTACCATTTCCTTCTGGAGCCCGGAGTATCCGCCATCTATGAAATGAAC	2460
	Qy	2461	GTCTCTACCTCCCTACATTTCTGTGCGCAATGCCCTCAGAAGGTACTACCTGTAATGGCAC	2520
	Db	2461	GTCTCTACCTCCCTACATTTCTGTGCGCAATGCCCTCAGAAGGTACTACCTGTAATGGCAC	2520
	Qy	2521	TGGACCGTGGACTGGCCCGGCCCGGTACNAATTTTCGGGCACCTACTTTTCGACTACAGACGG	2580
	Db	2521	TGGACCGTGGACTGGCCCGGCCCGGTACNAATTTTCGGGCACCTACTTTTCGACTACAGACGG	2580
	Qy	2581	TCCCTATAATGAGCCCGGAGAACCTTAATCGCTACTTGGACCACCAACGAGACACTGTATGTG	2640
	Db	2581	TCCCTATAATGAGCCCGGAGAACCTTAATCGCTACTTGGACCACCAACGAGACACTGTATGTG	2640
	Qy	2641	GAGCTCGTGTTCAGGGAGGAAACCCGGGTGTGCTGGGAATACCTCAGTCCGCTCGCTTG	2700
	Db	2641	GAGCTCGTGTTCAGGGAGGAAACCCGGGTGTGCTGGGAATACCTCAGTCCGCTCGCTTG	2700
	Qy	2701	GGGACCGAGAAGCAGGCCCTTGCCACGCCAGCTACACTTTGGGCCCATCGTGCCTCTGAG	2760
	Db	2701	GGGACCGAGAAGCAGGCCCTTGCCACGCCAGCTACACTTTGGGCCCATCGTGCCTCTGAG	2760
	Qy	2761	TGCTCCGTGTTCCTGCGAGGGGTAGTGGCCITTCAGTGCCTCTCTCGGAGGCGAGCATGT	2820
	Db	2761	TGCTCCGTGTTCCTGCGAGGGGTAGTGGCCITTCAGTGCCTCTCTCGGAGGCGAGCATGT	2820
	Qy	2821	CAGCCTTCAGCCACTGGGTACATTGGACTGGCCCTTTCTTGAATCTTAATGAGCAGCCCCG	2880
	Db	2821	CAGCCTTCAGCCACTGGGTACATTGGACTGGCCCTTTCTTGAATCTTAATGAGCAGCCCCG	2880
	Qy	2881	GGCTTCTCCCTGCCA	2895
	Db	2881	GGCTTCTCCCTGCCA	2895

DEPT. T. 2

RESULI 2
AR559323

ABS59323
ID ABS59323 standard: DNA: 2997 BP.

XX

AC ABS59323;

XX

DT 05-NOV-2002 (first entry)

Human zinc metalloprotease-like gene #1.

Human; NOX; cardiomyopathy; atherosclerosis; cell signal processing; breast cancer; Alzheimer's disease; epilepsy; Huntington's disease; anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis; neurodegeneration; Parkinson's disease; pain; stroke; endometriosis; autoimmune disease; allergy; addiction; asthma; transplantation; graft versus host disease; systemic lupus erythematosus; scleroderma; psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus; atherosclerosis; cholestasis; rheumatoid arthritis; diabetes; pancreatic thrombocytopenia; bleeding disorder; metabolic disorder; obesity; glucose transport defect; glomerulonephritis; hypercalcaemia; polycystic kidney disease; renal tubular acidosis; skin disorder;

QY 2161 CTCTACTGTATCGCAGAGGATTTGATTTCTTCTTTTCTTTGTCAAATAAAGTCAAAGAT 2220

KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
 KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
 KW haematopoietic disorder; urinary system disorder; osteoporosis; ds;
 KW dental disease; dental infection; growth disorder; reproductive disorder;
 KW hypogonadism; fertility disorder; viral infection; bacterial infection;
 KW parastic infection; metabolic pathway modulation; gene therapy; gene;
 KW zinc metalloproteinase; ADAM-TS 7; alpha-2-macroglobulin precursor;
 KW ileal sodium/bile acid cotransporter; probiotin; MT, CIP4; spinesin;
 KW macrophage stimulating protein precursor; fatty acid-binding protein;
 KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease;
 KW single nucleotide polymorphism; SNP.

OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT variation replace (472,G)
 FT /*tag= a
 FT variation replace (481,C)
 FT /*tag= b
 FT variation /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (1121,C)
 FT /*tag= c
 FT variation /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (1516,C)
 FT /*tag= d
 FT variation /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (1566,C)
 FT /*tag= e
 FT variation /standard_name= "Single nucleotide polymorphism (SNP)"

W0200233087-A2.

25-APR-2002.

17-OCT-2001; 2001WO-US32496.

17-OCT-2000; 2000US-241040P.

17-OCT-2000; 2000US-241058P.

17-OCT-2000; 2000US-241063P.

17-OCT-2000; 2000US-241243P.

20-OCT-2000; 2000US-242152P.

23-OCT-2000; 2000US-242482P.

23-OCT-2000; 2000US-242611P.

23-OCT-2000; 2000US-242612P.

24-OCT-2000; 2000US-242880P.

24-OCT-2000; 2000US-242881P.

29-DEC-2000; 2000US-259028P.

29-FEB-2001; 2001US-269813P.

25-APR-2001; 2001US-286324P.

29-MAY-2001; 2001US-294108P.

09-JUL-2001; 2001US-303698P.

16-OCT-2001; 2001US-0981151.

(CURA-) CURAGEN CORP.

Edinger S, Gerlach V, MacDougall JR, Malyankar UM, Smithson G;

Millet I, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shinkets RA;

Padigaruru M, Guo X, Patturajan M, Taupier RJ, Burgess CE;

Zerhusen BD, Kekuda R, Spytek KA, Gangolli EA, Fernandes ER;

Gorman L;

WP1; 2002-590434/63.
 P-PSDB; ABG76894.

Cyttoplasmic, nuclear, membrane bound and secreted polypeptides and

nucleic acids encoding the polypeptides for diagnosing and treating

e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and

diabetes

Claim 8; Page 12; 305pp; English.

The present invention relates to new NOVX (NOV1-10) polypeptides. The

CC molecules of the invention are useful for treating or preventing a
 CC NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or
 CC a disorder related to cell signal processing and metabolic pathway
 CC modulation in humans. NOVX polypeptides, nucleic acids and antibodies
 CC are useful for treating or preventing disorders or syndromes including
 CC breast cancer. Alzheimer's disease, epilepsy, Huntington's disease,
 CC anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis,
 CC neurodegeneration, Parkinson's disease, pain, stroke, autoimmune
 CC disease, allergies, addiction, asthma, endometriosis, graft versus host
 CC psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,
 CC atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes, obesity,
 CC thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 CC glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic
 CC kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 CC congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 CC muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 CC urinary system disorders, osteoporosis, dental disease and infection,
 CC growth and reproductive disorders, hypogonadism, fertility, and/or other
 CC pathologies and disorders, viral, bacterial, or parasitic infections.
 CC The present nucleic acid sequence encodes a NOVX protein of the
 CC invention.

XX
 SQ Sequence 2997 BP; 728 A; 783 C; 809 G; 673 T; 4 other;

Query Match 92.4%; Score 2675.4; DB 24; Length 2997;

Best Local Similarity 94.9%; Pred. No. 0;

Matches 2866; Conservative 0; Mismatches 5; Indels 150; Gaps 4;

QY 1 CGCTCTGATGAAGCCCGCGCGCGGATGAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 120

DB 1 CGCTCTGATGAAGCCCGCGCGCGGATGAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 60

QY 61 TTGGCGCAGTGGCGGAGCAGGTGAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 120

DB 61 TTGGCGCAGTGGCGGAGCAGGTGAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 120

QY 121 TCGGACAGCTGGAGCGAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 180

DB 121 TCGGACAGCTGGAGCGAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 180

QY 181 GCGATGTCGCGCGCTGTTTTCGCGAGCAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 240

DB 181 GCGATGTCGCGCGCTGTTTTCGCGAGCAGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 240

QY 241 GCGTGGAGCGCGAGCGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 300

DB 241 GCGTGGAGCGCGAGCGTCCCGGGCGCTCCACAGCGCGGAACCGCGGG 300

QY 301 AAAGGCGGACATGGATGAAGCTGGAACCATCGTTCTCAGCAAACTAACACGGAACA 360

DB 301 AAAGGCGGACATGGATGAAGCTGGAACCATCGTTCTCAGCAAACTAACACGGAACA 360

QY 361 GAAACCAAACTGATGTTCTCACTCAATATGACCTGGTCTCTGCTTACGAGGTGAC 420

DB 361 GAAACCAAACTGATGTTCTCACTCAATATGACCTGGTCTCTGCTTACGAGGTGAC 420

QY 421 CACAGGGCGATGATGTTCCCATGAATCATGACCATGACGGCGGAGAGACAGTGG 480

DB 421 CACAGGGCGATGATGTTCCCATGAATCATGACCATGACGGCGGAGAGACAGTGG 480

QY 481 GCGGTGTCGAGGTTGAGTCTCTTCACTCAATATGACCTGGTCTCTGCTTACGAGGTGAC 540

DB 481 GCGGTGTCGAGGTTGAGTCTCTTCACTCAATATGACCTGGTCTCTGCTTACGAGGTGAC 540

QY 541 ATGGATCTGAGGACTTCCAGGAGCCTAGTGGCTCTCTGCTTATTTGTCAGACGTTGGGA 600

DB 541 ATGGATCTGAGGACTTCCAGGAGCCTAGTGGCTCTCTGCTTATTTGTCAGACGTTGGGA 600

QY 601 AAGCAGGACCTAAGTCTGTGCGAGCTTTACCGCAGGAGGACTTCTGTTTCTATCAAGGC 660

DB 601 AAGCAGGACCTAAGTCTGTGCGAGCTTTACCGCAGGAGGACTTCTGTTTCTATCAAGGC 660

Db 2797 CGCTTGGGACCGAGAGAGCCCTCCAGCCAGCTACACTGGGCGCATCGTGGC 2856
 QY 2755 TCTGAGTGTCTCGTGTCTCTCGGAGGGGTAGTGCCTTCCAGTGCCTCTCTGGAGGCA 2814
 Db 2857 TCTGAGTGTCTCGTGTCTCTCGGAGGGGTAGTGCCTTCCAGTGTCTCTCTGGAGGCA 2916
 QY 2815 GCATGTGAGCTTCCAGCACTCGGTACATTTGCACTGCGCTTCTTGAATCCTTAATGAGCA 2874
 Db 2917 GCATGTGAGCTTCCAGCACTCGGTACATTTGCACTGCGCTTCTTGAATCCTTAATGAGCA 2976
 QY 2875 GCGCGGGCTTCTCCCTGGCA 2895
 Db 2977 GCGCGNGCTTCTCCCTGGCA 2997

RESULT 3

AAS97174
 ID AAS97174 standard; cDNA; 3675 BP.

XX AC AAS97174;

XX DT 26-FEB-2002 (first entry)

XX DE Human metalloprotease partial DNA sequence #3.

XX KW Human; protease; PCR primer; cytosolic; immunomodulator; cardiant;
 vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;
 hypertensive; hypoglycemic; neuroleptic; neuroprotective; anabolic;
 anorectic; antiinflammatory; aspartyl protease; cysteine protease;
 metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
 lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
 immune-related disease; cardiovascular disease; neuronal disease;
 migraine; sexual dysfunction; mood disorder; attention disorder;
 cognition disorder; hypotension; hypertension; psychotic disorder;
 dyskinesia; metabolic disorder; inflammatory disorder; ss.

XX OS Homo sapiens.

XX PN WO200183782-A2.

XX PD 08-NOV-2001.

XX PF 04-MAY-2001; 2001WO-US14431.

XX PR 04-MAY-2000; 2000US-201879P.

XX PA (SUGS-) SUGEN INC.

XX PI Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;

XX PI Payne V;

XX DR WPI; 2002-041502/05.

XX DR P-PSDB; AAU72891.

XX PT Novel protease polypeptide useful for screening for substances that may
 be used to treat, e.g., cancers, immune-related diseases,
 cardiovascular disease, migraine, pain, psychotic and inflammatory
 disorders.

XX PS Claim 30; Figure 1H; 232pp; English.

XX CC The invention relates to an isolated, enriched, or purified protease
 polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
 screen for substances (S) that may modulate its activity. Administering
 S (which modulates protease activity in vitro) may be used to treat a
 disease or disorder selected from cancers (e.g., of tissues, of blood or
 haematopoietic origin, of the breast, colon, lung, prostate, cervical,
 brain, ovarian, bladder or kidney), immune-related diseases and
 disorders, cardiovascular disease, brain or neuronal-associated diseases
 (e.g., central or peripheral nervous system diseases, migraine, pain,
 sexual dysfunction, mood disorders, attention disorders, cognition
 disorders, hypotension, hypertension, psychotic disorders, neurological
 disorders and dyskinesias), metabolic disorders and inflammatory

CC disorders. (I) may also be useful as a diagnostic tool for a disease or
 disorder such as those above. AAS97159-AAS97195 represent human
 CC protease coding sequences and primers of the invention.

XX Sequence 3675 BP; 874 A; 1009 C; 1019 G; 773 T; 0 other;

SQ Query Match 64.9%; Score 1877.8; DB 24; Length 3675;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 2289; Conservative 0; Mismatches 177; Indels 245; Gaps 6;

QY 202 GCAGGCACCTGCTGCGCCATGGACCCCGCAGCGCTGGAGAGCCCGAGCGTCC 261
 Db 69 GCAGGCACCTGCTGCGCCATGGACCCCGCAGCGCTGGAGAGCCCGAGCGTCC 128
 QY 262 GCGTCTCTCTCCACCGCGGAGCGCGCGCTGGATGGAAAAAGGCGGACATGATGAA 321
 Db 129 GCGTCTCTCTCCACCGCGGAGCGCGCGCTGGATGGAAAAAGGCGG----- 175
 QY 322 GCTGGAAACCATCGTTCTCAGCAAACTAACACAGGACAGAAAAACCAACACTGCATGTT 381
 Db 176 ----- 175
 QY 382 CTCACCTCAATATGACCTGCTCTGCTACGAGGTTGACCACAGGGCGGATACGTTGCC 441
 Db 176 -----AATATGACCTGCTCTGCTACGAGGTTGACCACAGGGCGGATACGTTGCC 228
 QY 442 CATGAAATCATACCATCATGCGCGGAGAGAGAGAGTGGCGTTCGAGGTTGAGTCT 501
 Db 229 CATGAAATCATACCATCATGCGCGGAGAGAGAGAGTGGCGTTCGAGGTTGAGTCT 288
 QY 502 CTTACCTTCGGCTCAAAAGGCCCGGAGGACGACTCCACATGGATCTGAGGACTTCCAGC 561
 Db 289 CTTACCTTCGGCTCAAAAGGCCCGGAGGACGACTCCACATGGATCTGAGGACTTCCAGC 348
 QY 562 AGCCTAGTGGCTCTCTGGCTTTATTGTGACAGAGCTTGGGAAAGACAGGCACTAACTGTG 621
 Db 349 AGCCTAGTGGCTCTCTGGCTTTATTGTGACAGAGCTTGGGAAAGACAGGCACTAACTGTG 408
 QY 622 CAGACTTTACCGCCAGAGAGACTTCTGTTTCTATCAAGGCTCTTTGCGATCAACAGAAAC 681
 Db 409 CAGACTTTACCGCCAGAGAGACTTCTGTTTCTATCAAGGCTCTTTGCGATCAACAGAAAC 468
 QY 682 TCGCCATCGCATGGAGGAAAGTTCTGTGAGGSGCTCCACTCGACACTCGAGCTCTGCAAC 741
 Db 469 TCTCTA-----GTGGCCCTTTCACTGCGCAAGGCTTGCA 504
 QY 742 AGTCAGAAATGTCCTCCGAGAGAGTGTGACTTCCGTGTGCTGCTGAGTGTCCGAGCAAC 801
 Db 505 GGCATGATACGAAACAGAGAGAGGAGGAGTACTTCTTAAGGCCACTTCTCTTCACTCTCA 564
 QY 802 AGCAGACGATTCAGAGGGGCGGCACTACAAGTGG--AAGCCTTACACTCAAGTAGAAGCG 859
 Db 565 TGGAAACTCGGAGAGCTGCCCAAGGCGAGCTGGCCATCCCACTACTCTACAGAGATCC 624
 QY 860 ACTTATGCAAACTCTACTGTATCGCAGAGAGATTTGATTTCTTCTTTCTTTGTCATAATA 919
 Db 625 ACAGAGCCCATGCTCTCTGGGCGAGTGAAGTCTCTGTGACCTC-----AAGG 672
 QY 920 AAGTCAAGATGGGACTCCATGCTCGGAGAGTAGCCGTAATGTTGTATAGATGGGATAT 979
 Db 673 ACATGGGAGCTGGCAGATCAACCCCTGCACAGCAGGACCTTCCGCTGGGACTGCCCAA 732
 QY 980 GTGAGCTCAGTGTGGTGTCCACATCTCG--GCACATGCGCCAGCTCCCAAGGAGACCTC 1038
 Db 733 AAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCTCCCAAGCTCCCAAGGAGACCTC 792
 QY 1039 TTCATCTTCCAGATGAGTATAAGTCTTGCTTACGGATAGCGCTCTCTTCTCAGGTC 1098
 Db 793 TTCATCTTCCAGATGAGTATAAGTCTTGCTTACGGATAGCGCTCTCTTCTCAGGTC 852
 QY 1099 CATGAAATGAAGAACTGAACGTCGAGACCTTGTGTGTGTCGACAAAAAGATGATGAA 1158
 Db 853 CATGAAATGAAGAACTGAACGTCGAGACCTTGTGTGTGTCGACAAAAAGATGATGAA 912

PP 11-OCT-2001; 2001WO-JP08913.
 XX
 PR 11-OCT-2000; 2000JP-0311309.
 PR 02-APR-2001; 2001JP-0102905.
 XX
 PA (KAZU-) KAZUSA DNA RES INST FOUND.
 PA (MTS-) MITSUBISHI PHARMA CORP.
 XX
 PI Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 PI Kanzaki K;
 XX
 DR WPI; 2002-372277/40.
 XX
 PT Human brain-originated ADAMTS family polypeptide and encoded gene,
 PT applicable in diagnosis and screening compounds for drug compositions
 PT in treating diseases due to e.g. neovascularisation
 XX
 XX Example 20; Page 155-159; 172pp; Japanese.
 XX
 CC The invention relates to a polypeptide belonging to the ADAMTS family is
 CC selected from sequences appearing as AAU79496, AAU79497 and AAU79499,
 CC a protein that contains the polypeptide, a protein having not less than
 CC 50% homology with the amino acid sequence of the polypeptides or a
 CC polypeptide modified from any of the polypeptides but with some amino
 CC acids deleted, substituted, added or inserted. Also included are the
 CC polynucleotides encoding the polypeptides (or their complementary
 CC strands or variants), a recombinant vector containing any of the
 CC polynucleotides, a transformant which is transformed with the recombinant
 CC vector, producing the polypeptide, protein or peptide by culturing the
 CC transformant, an antibody that can recognize the polypeptide, protein or
 CC peptide and screening compounds to promote or inhibit activity of the
 CC polypeptide or protein, or to promote or inhibit expression of the
 CC polynucleotide by using the polypeptide, protein, peptide, the
 CC polynucleotide, vector, transformant or/and antibody, particularly in
 CC the presence of a test compound for contact before evaluating the
 CC activity by measuring signal changes. The polypeptide and encoded gene
 CC are applicable in diagnosis and screening compounds for drug compositions
 CC in treating diseases due to neovascularisation, diabetic omentopathy,
 CC chronic rheumatoid arthritis, angioma, refractory skin and gastric ulcers
 CC and post-operative healing failure, including gene therapy.
 CC The gene encoding such polypeptide has conserved repolysin-type ZN-
 CC metalloprotease domain, disintegrin-like domain and TSR1 (chromospondin
 CC type 1) domain. Its encoded protein is characterised by high expression
 CC in ovaries, changes in expression dose depending on the sexual cycle, a
 CC decrease in tumour cell and location of the gene on the 5p-syntome
 CC deletion site on chromosome 5p15.2-15.3 (associated with
 CC Cri-du-chat syndrome). The present sequence is the expression Plasmid
 CC pFastBac1-HT-PJ01256 used to express DNA encoding an ADAMTS protein.
 XX
 SQ Sequence 8435 BP; 2072 A; 2174 C; 2149 G; 2040 T; 0 other;
 Query Match 64.9%; Score 1877.8; DB 24; Length 8435;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 2289; Conservative 0; Mismatches 177; Indels 245; Gaps 6;
 QY 202 GCAGGCACTGCGTGGCCCATGGACCCGCGAGCGGAGCGCCCTGGAGCCCGAGGTCCTCC 261
 DB 4120 GCAGGCACTGCGTGGCCCATGGACCCGCGAGCGGAGCGCCCTGGAGCCCGAGGTCCTCC 4179
 QY 262 GCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAGGCGGCGGACATGGATGAA 321
 DB 4180 GCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAGGCGGCGGACATGGATGAA 4226
 QY 322 GCTGGAACCATGCTTCTCAGCAACCTAACACAGGAGAACAAACCAACACTGCATGTT 381
 DB 4227 ----- 4226
 QY 382 CTCACCTCAATGACCTGGTCTCTGCTCTACGAGTTGACCAAGGCGGATACGTGTCC 441
 DB 4227 -----AATATGACCTGGTCTCTGCTCTACGAGTTGACCAAGGCGGATACGTGTCC 4279
 QY 442 CATGAATCATGCAACATCAGCGCGGAGAGCAAGTGCCTGTCCGAGGTGATGCT 501
 DB 4427 -----

4280 CATGAATCATGCAACATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGTTGAGTCT 4339
 QY 502 CTTACACCTTCGGCTGAAAGGCCCGCCAGGCACGACTTCCACATGATCTGAGGACTTCCACG 561
 DB 4340 CTTACACCTTCGGCTGAAAGGCCCGCCAGGCACGACTTCCACATGATCTGAGGACTTCCACG 4399
 QY 562 AGCCTAGTGGCTCCTCGCTTTATTTGTCAGAGCTTGGGAAAGACAGGCACTAAGTCTGTG 621
 DB 4400 AGCCTAGTGGCTCCTCGCTTTATTTGTCAGAGCTTGGGAAAGACAGGCACTAAGTCTGTG 4459
 QY 622 CAGACTTTACCGCCAGAGGACTTCTCTTTCTATCAAGGCTCTTTGGGATCACACAGAAC 681
 DB 4460 CAGACTTTACCGCCAGAGGACTTCTCTTTCTATCAAGGCTCTTTGGGATCACACAGAAC 4519
 QY 682 TCGCATTCGCATCGAGGAGGAGTTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTTCGCAAC 741
 DB 4520 TCCTCA-----GTGGCCCTTTCAACCTGCCAGGCTGTGCA 4555
 QY 742 AGTCAGAAATGTCCTCCGGGACAGTGTGACTTCCGTGCTGCTCAGTGTCCGAGGACACAC 801
 DB 4556 GGCATGATGAAACAGAGAGGAGGAGGATTTCTTCTTAAGGCCACTTCCCTTCACACCTCTCA 4615
 QY 802 AGCAGACGATTTCAGAGGGCGGCACTACAAAGTGG--AAGCCTTACACTCAAGTAGAAGCCG 859
 DB 4616 TGGAACTCGGAGAGCTGCCAAGGAGCTGCCATCCACGCTACTGTACAGAGATCC 4675
 QY 860 ACTTATGCAAACTCTACTGTATCGCAAGAGATTTGATTTCTTTCTTTTCTTTTCTTTTCTT 919
 DB 4676 ACAGAGCCCATGCTCTCTGGGGCAGTGAGGCTCTGTGACCTC-----AAGG 4723
 QY 920 AAGTCAAGATGGAGCTCCATGCTCGAGGATAGCGTAATGTTTGTATAGATGGATAT 979
 DB 4724 ACATGGAGCTGGCAATCAACCCCTGCACAGCAGGAGCTTCGCTGGGACTGCCACAA 4783
 QY 980 GTGAGCTCAGTGTGGTTCACATCTGC--GCACATGCCCGCAGCCTCCCAAGGAGACCTC 1038
 DB 4784 AAGCAGCATTTCTGTGAAGACGCAAGAAATACATGCCCGCAGCTCCCAAGGAGACCTC 4843
 QY 1039 TTCACTTTGCCAGATGATAGTCTTGTCTTACGCAAGAGGCTCTCTCTGAGAGTCC 1098
 DB 4844 TTCACTTTGCCAGATGATAGTCTTGTCTTACGCAAGAGGCTCTCTCTGAGAGTCC 4903
 QY 1099 CATGAATGAAGAACTGAACCTGGAGAACCTTGGTGGTGTGCAAAAAAGATGATGCAA 1158
 DB 4904 CATGAATGAAGAACTGAACCTGGAGAACCTTGGTGGTGTGCAAAAAAGATGATGCAA 4963
 QY 1159 AACCATGGCCATGAATATATCACCCTACGCTCTCAGCATCTCAGATGGTATCTGCT 1218
 DB 4964 AACCATGGCCATGAATATATCACCCTACGCTCTCAGCATCTCAGATGGTATCTGCT 5023
 QY 1219 TTATTTAAAGA----- 1229
 DB 5024 TTATTTAAAGATGGAACATAGGAGGAAACATCAAGATTCGATTCGATTCCT 5083
 QY 1230 ----- 1229
 DB 5084 CTAGAGATGAACAGCCAGGACTGGTATAGTCAACCGCAGACACACCTTTAAGTAGC 5143
 QY 1230 -----TGATTTGATGGGAAAGATGGGACTCGTCAATGACACCGCATC 1272
 DB 5144 TTCTGCCAGTGGCAGTCTGGATTTGATGGGAAAGATGGGACTCGTCAATGACACCGCATC 5203
 QY 1273 TTACTGATCGGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGACATTTGGGATTT 1332
 DB 5204 TTACTGATCGGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGACATTTGGGATTT 5263
 QY 1333 GCACCCATAAGTGAATGTGTAGTAATATCGCAGCTGCACGATTAATGAAGATACAGT 1392
 DB 5264 GCACCCATAAGTGAATGTGTAGTAATATCGCAGCTGCACGATTAATGAAGATACAGT 5323
 QY 1393 CTTGGACTGGCCTTACCATTTGCCATTTGAGTCTGGAACACACTTTGGCATGATCATGAT 1452
 DB 5324 CTTGGACTGGCCTTACCATTTGCCATTTGAGTCTGGAACACACTTTGGCATGATCATGAT 5383


```

DB 5754 CGTATGGAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAAAGCACAATTTGTGGG 5813
QY 1813 CATGACATGTGTGCGGGAGGACAGTGTGTGAATATGTGATGAGGCCCAAGCCC 1872
DB 5814 CATGACATGTGTGCGGGAGGACAGTGTGTGAATATGTGATGAGGCCCAAGCCC 5873
QY 1873 ACCCATGGCCACTGTGCGGACTGTGCTTTCTTGGTCCCATCTCTCAGGACCTGCGGAGG 1932
DB 5874 ACCCATGGCCACTGTGCGGACTGTGCTTTCTTGGTCCCATCTCTCAGGACCTGCGGAGG 5933
QY 1933 GAGATATCTATAGAGTGGCTCTGACCAACCCAGCCATCGCATGAGGAGATTC 1992
DB 5934 GAGATATCTATAGAGTGGCTCTGACCAACCCAGCCATCGCATGAGGAGATTC 5993
QY 1993 TGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCTGCAAAATGTCCCGGGACAGT 2052
DB 5994 TGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCTGCAAAATGTCCCGGGACAGT 6053
QY 2053 GTTGACTTCGCTGCTGCTCAGTGTGCGGAGGACACACAGACAGATTCAGAGCGGCGAC 2112
DB 6054 GTTGACTTCGCTGCTGCTCAGTGTGCGGAGGACACACAGACAGATTCAGAGCGGCGAC 6113
QY 2113 TACAAGTGAAGCCTTACACTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 2172
DB 6114 TACAAGTGAAGCCTTACACTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 6173
QY 2173 GCGAAGGATTTGATTTCTCTTTCTTTCTGCAATTAAGTCAAGATGAGGACCTCCATGC 2232
DB 6174 GCGAAGGATTTGATTTCTCTTTCTTTCTGCAATTAAGTCAAGATGAGGACCTCCATGC 6233
QY 2233 TCGGAGCATAGCGGTAAATGTTGTATAGATGGATGTGAGAGATTTGGATGTGCAAT 2292
DB 6234 TCGGAGCATAGCGGTAAATGTTGTATAGATGGATGTGAGAGATTTGGATGTGCAAT 6293
QY 2293 GTCTTTGATCTGATCTCTGAGAGCTCTGTGGGGTGTGTAACGGGAATTAACCTAGCC 2352
DB 6294 GTCTTTGATCTGATCTCTGAGAGCTCTGTGGGGTGTGTAACGGGAATTAACCTAGCC 6353
QY 2353 TGCAAGATTCACAGGGGTCTCTACACCAAGCACCAACCAACAGTATTAACATG 2412
DB 6354 TGCAAGATTCACAGGGGTCTCTACACCAAGCACCAACCAACAGTATTAACATG 6413
QY 2413 GTCAAGATTCCTTCTGGAGCCGGAGTATCCGATCTATGAATGAAGCTCTACCTCC 2472
DB 6414 GTCAAGATTCCTTCTGGAGCCGGAGTATCCGATCTATGAATGAAGCTCTACCTCC 6473
QY 2473 TACATTTCTGTGGCAATGCCCTCAGAGGTACTACTGGAATGGGCACTGGACCGTGGAC 2532
DB 6474 TACATTTCTGTGGCAATGCCCTCAGAGGTACTACTGGAATGGGCACTGGACCGTGGAC 6533
QY 2533 TGGCCGGCGGCTACAAATTTTGGGCACTACTTTGCACTACAGAGGTCTCTATATGAG 2592
DB 6534 TGGCCGGCGGCTACAAATTTTGGGCACTACTTTGCACTACAGAGGTCTCTATATGAG 6593
QY 2593 CCGGAGACTTAATCTCTACTGGAACCAACCAACAGACACTGATTTGGAGCTCTGTTT 2652
DB 6594 CCGGAGACTTAATCTCTACTGGAACCAACCAACAGACACTGATTTGGAGCTCTGTTT 6653
QY 2653 CAGGGAAGAAACCCGGGTGTGCTTGGGAATATCTCCATGCCCTTGGGACCGGAGAG 2712
DB 6654 CAGGGAAGAAACCCGGGTGTGCTTGGGAATATCTCCATGCCCTTGGGACCGGAGAG 6713
QY 2713 CAGCCCGCTCCAGCCAGCTACACTTGGGCACTCTGCGCTCTGAGTGTCCGTGTC 2772
DB 6714 CAGCCCGCTCCAGCCAGCTACACTTGGGCACTCTGCGCTCTGAGTGTCCGTGTC 6773
QY 2773 TCGGAGAGGGG 2783
DB 6774 TCGGAGAGGGG 6784

```

RESULT 7

```

AAL43654
ID AAL43654 standard; DNA; 3675 BP.
XX
AC AAL43654;
XX
DT 05-SEP-2002 (first entry)
XX
DE Human MDTs9 protease coding sequence.
XX
KW Human; gene; ds; MDTs9 protease; TGF-beta inhibitor;
transforming growth factor-beta inhibitor; chronic renal failure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3675
FT /tag= a
FT /product= "Human MDTs9 protease"
XX
XX MO200251998-A1.
XX
XX 04-JUL-2002.
XX
XX 21-DEC-2001; 2001WO-JP11251.
XX
XX 25-DEC-2000; 2000JP-0393372.
XX
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
XX
XX Yamaji N, Nishimura K, Abe K, Ogino M;
XX
XX WPI; 2002-508888/54.
XX
XX P-PSDB; AA015254.
XX
XX Protease MDTs9 and encoded polynucleotide, applicable in diagnosis and
screening TGF-beta inhibitors for treatment of chronic renal failure -
XX
XX Example 2; Page 41-48; 60pp; Japanese.
XX
XX The invention comprises the amino acid and coding sequence of the human
MDTs9 protease. The MDTs9 DNA and protein sequences of the invention are
useful for screening for transforming growth factor (TGF)-beta inhibitors
and for the treatment of chronic renal failure. The present DNA sequence
encodes the human MDTs9 protease of the invention.
XX
SQ Sequence 3675 BP; 873 A; 1007 C; 1020 G; 775 T; 0 other;
Query Match 64.8%; Score 1874.6; DB 24; Length 3675;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;
QY 202 GCAGGACCTGCTGCGGCATGGGACCCGCGAGCGCTGGAGCCCGAGCGTCCC 261
DB 69 GCAGGACCTGCTGCGGCATGGGACCCGCGAGCGCTGGAGCCCGAGCGTCCC 128
QY 262 GCGTCTCTCTCCACCGCGGAGCGCGGCTGGATGAAAGGCGGACATGGATGA 321
DB 129 GCGTCTCTCTCCACCGCGGAGCGCGGCTGGATGAAAGGCGG 175
QY 322 GCTGGAACCATCGTTCGCAAACTAACACAGAAACAGAAACACATCGCATGTT 381
DB 176 ----- 175
QY 382 CTCACCTCAATATGACCTGCTCTCCCTACGAGTTGACACAGGCGGATTCGTGTC 441
DB 176 -----AATATGACCTGCTCTCTCCCTACGAGTTGACACAGGCGGATTCGTGTC 228
QY 442 CATGAATCATGCACCATCAGCGCGGAGAGAGAGTGGCCGTGTCCGAGGTGAGTCT 501
DB 229 CATGAATCATGCACCATCAGCGCGGAGAGAGAGTGGCCGTGTCCGAGGTGAGTCT 288
QY 502 CTTCACTCTCGCTGAGAGGCCCGGAGCCAGCTTCCACATGATCTGAGACTTCCAGC 561
DB 502 CTTCACTCTCGCTGAGAGGCCCGGAGCCAGCTTCCACATGATCTGAGACTTCCAGC 561

```


Db 289 CTTCACTTCGGCTGAAGAGCTCCAGGCAAGACTTCCAGTGTGATCTGAGGACTTCCAGC 348
Qy 562 AGCTAGTGGCTCTGGCTTTATTTGTCAGACGCTTGGGAAAGACAGGCACTAAGTCTGTG 621
Db 349 AGCTAGTGGCTCTGGCTTTATTTGTCAGACGCTTGGGAAAGACAGGCACTAAGTCTGTG 408
Qy 622 CAGACTTTACCGCCAGAGGACTCTGTTCTATCAAGGCTCTTTGGGATCACAAGAAAC 681
Db 409 CAGACTTTACCGCCAGAGGACTCTGTTCTATCAAGGCTCTTTGGGATCACAAGAAAC 468
Qy 682 TGGCCATCGCATGGAGGAGTCTGTGAGGAGCTCCAGCTCGGACTCTGAAGCTCTCAAC 741
Db 469 TCTCTCA-----GTGGCCCTTTCAACCTGCGCAAGGCTTGTCA 504
Qy 742 AGTCAGAAATGTCCCGGACAGTGTGACTTCCGTCTGCTCAGTGTGCGGAGCACAAC 801
Db 505 GGCATGNTAGCAACAGAGAGGAGATTACTTCTTAAGGCGACTTCTTCAACACTCTCA 564
Qy 802 AGCAGACGATTCAAGAGGCGGCACACTACAAGTGG--AAGCCTTACACTCAAGTGAAGGCG 859
Db 565 TGGAAACTCGGAGAGCTGCCAAGGCGAGCTCGGCATCCCACTGCTGTACAAGAGATCC 624
Qy 860 ACTTATGCAACTCTACTGTATCCAGAGGATTTGATTTCTTCTTCTTCTTCTCAATA 919
Db 625 ACAGAGCCCCATGCTCTTGGGCGCAGTGTGAGTCTGTGACCTC-----AAG 672
Qy 920 AAGTCAAGATGGGACTCCATGCTCGGAGGATGCCGTAATGTTGTATAGTGGATAT 979
Db 673 ACATGGGAGCTGGCAACATCAACCCCTGCACAGAGCGACCTTGGCCTGGGACTGCCACA 732
Qy 980 GTGAGCTCAGTGTGGTGTCCACATCTGC-GCACATGCCCGAGCTCCCAAGGAGACCTC 1038
Db 733 AAGCAGCATTTCTGTGGAAGCGCAAGAAATACATGCCCGAGCTCCCAAGGAGACCTC 792
Qy 1039 TTCACTTGGCAGATGATATAAGTCTTGTCTTACGATTAAGCGCTCTCTTCTGAGTCC 1098
Db 793 TTCACTTGGCAGATGATATAAGTCTTGTCTTACGATTAAGCGCTCTCTTCTGAGTCC 852
Qy 1099 CATGAATGAAGACTGAAGCTGGAGACCTTGGTGGTGGTGGCAAAAGAGTGGCAA 1158
Db 853 CATGAATGAAGACTGAAGCTGGAGACCTTGGTGGTGGTGGTGGTGGCAAAAGAGTGGCAA 912
Qy 1159 AACATGGCCATGAAGAAATATACACACCTTACGCTCTCAGATCTCAACATGTTGTCT 1218
Db 913 AACATGGCCATGAAGAAATATACACACCTTACGCTCTCAGATCTCAACATGTTGTCT 972
Qy 1219 TTATTCNAGA----- 1229
Db 973 TTATTCNAGAATGGACATAGGAGGAACATCAACATTGCCAATGTAGTCTGATCTT 1032
Qy 1230 ----- 1229
Db 1033 CTAGAAGATGAACAGCCAGGACTGGTGATAGTCAAGTCAACAGCAGACCAACCTTAAAGTAC 1092
Qy 1230 -----TGGATTGATGGGAAAGATGGGACTCTGTCATGACCAAGCCTC 1272
Db 1093 TTCTCCAGTGGCACTCTGGATTGATGGGAAAGATGGGACTCTGTCATGACCAAGCCTC 1152
Qy 1273 TTACTGACTGGTCTGGATATATGTTCTGGGAAAGATGAGCCCTGTGACACTTTGGGATTT 1332
Db 1153 TTACTGACTGGTCTGGATATATGTTCTGGGAAAGATGAGCCCTGTGACACTTTGGGATTT 1212
Qy 1333 GCACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGCAGATTAATGAGATACAGGT 1392
Db 1213 GCACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGCAGATTAATGAGATACAGGT 1272
Qy 1393 CTTGAGCTGGCTTCCACATTTGCCCATGAGTCTGGACAAACTTTGGCATGATTCATGAT 1452
Db 1273 CTTGAGCTGGCTTCCACATTTGCCCATGAGTCTGGACAAACTTTGGCATGATTCATGAT 1332
Qy 1453 GGAGAGGGAACATGTGTAAGAAAGTCCGAGGCGCAACATCATGTCCTTACATTTGGCAGGA 1512
Db 1333 GGAGAGGGAACATGTGTAAGAAAGTCCGAGGCGCAACATCATGTCCTTACATTTGGCAGGA 1392

Qy 1513 CGCAATGAGTCTTCTCTGCTCACCTCGAGCCGAGCTATCTACACAATTTCTAAGC 1572
Db 1393 CGCAATGAGTCTTCTCTGCTCACCTCGAGCCGAGTATCTACACAATTTCTAAGC 1452
Qy 1573 ACCGCTCAAGCTATCTGCTGCTGATCAGCCAAAGCCTGTGAAGAAATACAAGTATCT 1632
Db 1453 ACCGCTCAAGCTATCTGCTGCTGATCAGCCAAAGCCTGTGAAGAAATACAAGTATCT 1512
Qy 1633 GAGAAATGCGAGGAAATATATGATGCAACACACAGCTGCAAGTGGCAGTTCGGAG 1692
Db 1513 GAGAAATGCGAGGAAATATATGATGCAACACACAGCTGCAAGTGGCAGTTCGGAG 1572
Qy 1693 AAAGCCAAAGCTCTGCAATGCTGAGCTTTAAAGAGACATCTGTAAGCCCTGTGGTGCAT 1752
Db 1573 AAAGCCAAAGCTCTGCAATGCTGAGCTTTAAAGAGACATCTGTAAGCCCTGTGGTGCAT 1632
Qy 1753 GGTATTGGAAGAAATGTGAGACTAAATTTATGCCAGAGCAGAGAGGCAAAATTTGTGG 1812
Db 1633 GGTATTGGAAGAAATGTGAGACTAAATTTATGCCAGAGCAGAGAGGCAAAATTTGTGG 1692
Qy 1813 CATCAGATGTGTGCCCGGAGGAGCAGTGTGTGAATATGTTGATGATGAGAGGCAAGCC 1872
Db 1693 CATCAGATGTGTGCCCGGAGGAGCAGTGTGTGAATATGTTGATGATGAGAGGCAAGCC 1752
Qy 1873 ACCCATGGCACTGTGCTGGA-CTGCTCTTGTGTGCCATGCTCCAGGACCTGCGGAGG 1932
Db 1753 ACCCATGGCACTGTGCTGGA-CTGCTCTTGTGTGCCATGCTCCAGGACCTGCGGAGG 1812
Qy 1933 GGAGTATCTATAGGAGTCTGCTGCAACCAAGCCCAAGCCATCGCATCGAGGAGGATTC 1992
Db 1813 GGAGTATCTATAGGAGTCTGCTGCAACCAAGCCCAAGCCATCGCATCGAGGAGGATTC 1872
Qy 1993 TGTGAGGCTCTCACTGCGACTCTGGAAGCTCTGCAAGCTCAGATCAGAAATGTCGCGGAG 2052
Db 1873 TGTGAGGCTCTCACTGCGACTCTGGAAGCTCTGCAAGCTCAGATCAGAAATGTCGCGGAG 1932
Qy 2053 GTTGACTTCCGTGCTGCTCAGTGTGCCGAGCACAACAGCAGACGATTCAGAGGCGGCGAC 2112
Db 1933 GTTGACTTCCGTGCTGCTCAGTGTGCCGAGCACAACAGCAGACGATTCAGAGGCGGCGAC 1992
Qy 2113 TACAAGTGGAAAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTTACTGTATC 2172
Db 1993 TACAAGTGGAAAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTTACTGTATC 2052
Qy 2173 GCAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2232
Db 2053 GCAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2112
Qy 2233 TCGAGAGATAGCCGATGTTGTATAGATGGGATATGTGAGAGAGTGGATGGATGGCAAT 2292
Db 2113 TCGAGAGATAGCCGATGTTGTATAGATGGGATATGTGAGAGAGTGGATGGATGGCAAT 2172
Qy 2293 GTCCTTGGATCTGATGCTGTGTAAGACCTCTGTGGGGTGTGTAACGGGAATAACTCAGCC 2352
Db 2173 GTCCTTGGATCTGATGCTGTGTAAGACCTCTGTGGGGTGTGTAACGGGAATAACTCAGCC 2232
Qy 2353 TGCAGGATTCAGGGGTCTCTACACAGACCAACACACACAGTATTTATCAGATG 2412
Db 2233 TGCAGGATTCAGGGGTCTCTACACAGACCAACACACACAGTATTTATCAGATG 2292
Qy 2413 GTCAACCATCTCTCGAGGCGCGGAGTATCCGANTCTATGAATGAACCTCTCTACCTCC 2472
Db 2293 GTCAACCATCTCTCGAGGCGCGGAGTATCCGANTCTATGAATGAACCTCTCTACCTCC 2352
Qy 2473 TACATTTCTGTGCGCAATGCTCTCAGAGGTA-CTACTGAATGGGCACTGGA-CCGTGGAC 2532
Db 2353 TACATTTCTGTGCGCAATGCTCTCAGAGGTA-CTACTGAATGGGCACTGGA-CCGTGGAC 2412
Qy 2533 TGGCCCGCGGCTACAAATTTTCGGGCACTACTTTTCGACTACAGAGCGTCTTATATGAG 2592
Db 2413 TGGCCCGCGGCTACAAATTTTCGGGCACTACTTTTCGACTACAGAGCGTCTTATATGAG 2472

Db 1033 CTAGAAGATGAA CAGGCGAGCTGGTGTAAGTACCCAGCAGACCAACCTTAACTAGTAC 1092
 Qy 1230 -----TGGATTGATGGGAAAGATGGGACTCGTCATGACCAAGCCATC 1272
 Db 1093 TTCTCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCATGACCAAGCCATC 1152
 Qy 1273 TTACTGACTCGTCTGGATATATGTTCTTGGAAAGATGGGACTCGTCATGACCAAGCCATC 1332
 Db 1153 TTACTGACTCGTCTGGATATATGTTCTTGGAAAGATGGGACTCGTCATGACCAAGCCATC 1212
 Qy 1333 GCACCCATAAGTGGATGTTAGTAAATATCGCAGCTGCAAGATTAATGAAGATACAGGT 1392
 Db 1213 GCACCCATAAGTGGATGTTAGTAAATATCGCAGCTGCAAGATTAATGAAGATACAGGT 1272
 Qy 1393 CTTGGACTCGCTTCACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
 Db 1273 CTTGGACTCGCTTCACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
 Qy 1453 GGAGAAGGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCCTTACATGCGAGGA 1512
 Db 1333 GGAGAAGGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCCTTACATGCGAGGA 1392
 Qy 1513 CGCAATGAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
 Db 1393 CGCAATGAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
 Qy 1573 ACCGCTCAAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
 Db 1453 ACCGCTCAAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
 Qy 1633 GAGAAATGCGCAGGAAATATATGATGCAACACACACAGTGGAGTGGCAGTTCGGAGAG 1692
 Db 1513 GAGAAATGCGCAGGAAATATATGATGCAACACACACAGTGGAGTGGCAGTTCGGAGAG 1572
 Qy 1693 AAAGCCAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
 Db 1573 AAAGCCAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
 Qy 1753 GGTATGGAAGGAAATGAGACTAAATTTATCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1812
 Db 1633 GGTATGGAAGGAAATGAGACTAAATTTATCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1692
 Qy 1813 CATGACATGTGCTCCGAGGAGCAGTGTGTAATATGCTGTAAGGAGCCCAAGGCC 1872
 Db 1693 CATGACATGTGCTCCGAGGAGCAGTGTGTAATATGCTGTAAGGAGCCCAAGGCC 1752
 Qy 1873 ACCCATGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932
 Db 1753 ACCCATGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1812
 Qy 1933 GGAGTATCTCATAGAGTGCCTCTGCAACCAACCAAGCCATTCGATGAGAGGAGTTC 1992
 Db 1813 GGAGTATCTCATAGAGTGCCTCTGCAACCAACCAAGCCATTCGATGAGAGGAGTTC 1872
 Qy 1993 TGTGAGGCTCCACTCGCAGCTGAGCTCTGCAACCAAGCAGTCAAGATGTCCTCGGAGCAGT 2052
 Db 1873 TGTGAGGCTCCACTCGCAGCTGAGCTCTGCAACCAAGCAGTCAAGATGTCCTCGGAGCAGT 1932
 Qy 2053 GTTGAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2112
 Db 1933 GTTGAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2052
 Qy 2113 TACAGTGGAGGCTTACCTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 2172
 Db 1993 TACAGTGGAGGCTTACCTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 2052
 Qy 2173 GCAGAAGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 2232
 Db 2053 GCAGAAGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 2112
 Qy 2233 TCGAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2292
 Db 2113 TCGAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2172

Qy 2293 GTCCCTTGGATCTGATGCTGTTGAAGACCTCTGTGGGGTGTGTAAACGGGAATTAACCTAGCC 2352
 Db 2173 GTCCCTTGGATCTGATGCTGTTGAAGACCTCTGTGGGGTGTGTAAACGGGAATTAACCTAGCC 2232
 Qy 2353 TGCAGATTTACAGGGGTCTCTTACACCAAGCACCACACACCAACCAAGTATTATCAGATG 2412
 Db 2233 TGCAGATTTACAGGGGTCTCTTACACCAAGCACCACACACCAACCAAGTATTATCAGATG 2292
 Qy 2413 GTCAACATTTCTTCTGGAGCCCGGAGTATCCGATCTATGAAATGAACGCTCTCTACCTCC 2472
 Db 2293 GTCAACATTTCTTCTGGAGCCCGGAGTATCCGATCTATGAAATGAACGCTCTCTACCTCC 2352
 Qy 2473 TACATTTCTGTCGCAATGCCCTCAGAAAGTACTACCTGAATGGCACTGGACCGTGGAC 2532
 Db 2353 TACATTTCTGTCGCAATGCCCTCAGAAAGTACTACCTGAATGGCACTGGACCGTGGAC 2412
 Qy 2533 TGGCCCGCCGGTACAAATTTTTCGGGCACTACTTTCGACTACAGACGGTCTCTAATAGAG 2592
 Db 2413 TGGCCCGCCGGTACAAATTTTTCGGGCACTACTTTCGACTACAGACGGTCTCTAATAGAG 2472
 Qy 2593 CCGGAGAACTTAATCGTACTGAGCAACCAACCAAGACACTGATTGTGGAGCTCTCTGTTT 2652
 Db 2473 CCGGAGAACTTAATCGTACTGAGCAACCAACCAAGACACTGATTGTGGAGCTCTCTGTTT 2532
 Qy 2653 CAGGGAAGGAAACCCGGGTGTTGCCCTGGGAAATATCTCCATGCTCGCTTGGGACCGAGAG 2712
 Db 2533 CAGGGAAGGAAACCCGGGTGTTGCCCTGGGAAATATCTCCATGCTCGCTTGGGACCGAGAG 2592
 Qy 2713 CAGCCCTCTGCCAGCCAGCTACACTTGGGCACTGCTGGGCTCTCTGAGTCTCGCTGCTCC 2772
 Db 2593 CAGCCCTCTGCCAGCCAGCTACACTTGGGCACTGCTGGGCTCTCTGAGTCTCGCTGCTCC 2652
 Qy 2773 TCGGAGGGGG 2783
 Db 2653 TCGGAGGGGG 2663

RESULT 9

ABS57768

ID ABS57768 standard; cDNA; 4042 BP.

XX ABS57768;

AC AC

XX DT

XX 05-FEB-2003 (first entry)

XX cDNA encoding novel human protease #2.

XX Human; protease; gene therapy; obesity; gene; ss.

XX Homo sapiens.

XX Key

XX CDS

XX Location/Qualifiers

XX 99..3773

XX /tag= a

XX /product= "Novel human protease"

XX US6448388-B1.

XX 10-SEP-2002.

XX 15-AUG-2001; 2001US-0930872.

XX 16-AUG-2000; 2000US-225852P.

XX (LEXI-) LEXICON GENETICS INC.

XX Fiddle CJ, Hilbun E;

XX WPI; 2003-074103/07.

XX P-PSDB; ABG72431.

XX New nucleic acid encoding novel human protein (NHP), useful for the

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT preparation of a medicament for diagnosing, preventing or treating
 PT disorders or diseases associated with the NHE, e.g., obesity -

PS Disclosure; Column 25-30; 17pp; English.

XX The invention describes a new isolated nucleic acid comprising a
 CC sequence that encodes a fully defined protein sequence comprising 491
 CC or 1224 amino acids or that hybridises under stringent conditions with
 CC the 3675-bp sequence or its complement. The nucleic acid is useful for
 CC the preparation of a medicament for diagnosing, preventing or treating
 CC disorders (e.g. using gene therapy) or diseases associated with the
 CC novel human proteins, e.g., obesity. This sequence encodes a novel
 CC human protease.

XX SQ Sequence 4042 BP; 947 A; 1117 C; 1130 G; 848 T; 0 other;

Query Match 64.8%; Score 1874.6; DB 25; Length 4042;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;

QY 202 GCAGGCACTCGTGGCCATGGGACCGCGAGCGGCGCTGGAGCCCGAGCGTCCC 261
 DB 167 GCAGGCACTCGTGGCCATGGGACCGCGAGCGGCGCTGGAGCCCGAGCGTCCC 226
 QY 262 GCGTCCCTCTCCACCGCGAGCGCGCGCTGGATGAAAGCGGACATGGATGAA 321
 DB 227 GCGTCCCTCTCCACCGCGAGCGCGCGCTGGATGAAAGCGGCGG----- 273
 QY 322 GCTGGAAACCATGCTTCAGCAAACTAACACAGGACAGAAACCAACACATGCGATGT 381
 DB 274 ----- 273
 QY 382 CTCACCTCAATATGACCTGCTCTGCTACGAGGTTGACACAGGCGGATACGTTGCC 441
 DB 274 -----AATATGACCTGCTCTGCTACGAGTTGACACAGGCGGATACGTTGCC 326
 QY 442 CATGAATCATGACCATCAGCGCGGAGAGAGCAGTGGCGCTGTCGAGGTTGAGTCT 501
 DB 327 CATGAATCATGACCATCAGCGCGGAGAGAGCAGTGGCGCTGTCGAGGTTGAGTCT 386
 QY 502 CTTACCTTCGGCTGAAAGCGCCCGAGCAGCTTCCACATGATCTGAGGACTTCCAGC 561
 DB 387 CTTACCTTCGGCTGAAAGCGCTCAGGACAGCTTCCACATGATCTGAGGACTTCCAGC 446
 QY 562 AGCTATGAGTCTCTGCTTATGTCAGACAGTGGGAAAGACAGGACATAGTCTGTG 621
 DB 447 AGCTATGAGTCTCTGCTTATGTCAGACAGTGGGAAAGACAGGACATAGTCTGTG 506
 QY 622 CAGACTTTACCGCCAGAGGACTCTGTTTCTATCAGGCTCTTTGGGATCACAAGAAAC 681
 DB 507 CAGACTTTACCGCCAGAGGACTCTGTTTCTATCAGGCTCTTTGGGATCACAAGAAAC 566
 QY 682 TCGCCATCGCATGAGGAGAGTCTGTGAGGCTCCACTCGCATCTGAGGCTCTGCAAC 741
 DB 567 TCCTCA -----GTGCGCTTTCAACCTTGCCAGGCTTTGTCA 602
 QY 742 AGTCAGAAATGTCGCGGACAGTGTGACTTTCGCTGCTGCTGAGTGGCGAGCACAAC 801
 DB 603 GGCATGATACGACAGAGAGGCGAGATTTCTCTTAGGCGCCTCTCTTCAACCTCTCA 662
 QY 802 AGCAGAGATTCAGAGGCGGCACTACAAGTGG--AAGCCTTACATCAAGTAGAAGCGG 859
 DB 663 TGGAACTCGGAGAGTGGCCAAAGGAGAGTGGCCATCCCAAGCTACTGTACAAAGAGATCC 722
 QY 860 ACTTATGCAAACTCTACTGATCCGAGAGAGATTTGATTTCTTTCTTTCTTCTCAATA 919
 DB 723 ACAGAGCCCATGCTCTCTGGGCGCAGTGAAGTCTCTGTGACCTC-----AAGG 770
 QY 920 AAGTCAAGATGGGACTCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATAT 979
 DB 771 ACATGGGAGCTGGACATCAACCCCTGCAAGCAGGACGACCTTCGCTGGGACTGCCACA 830
 QY 980 GTGAGCTCAGTGTGTGTCACATCTGC-GCACATGCCCGCAGCTCCCAAGGAGACCTC 1038

DB 831 AAGCAGCATTTCTGTGAGAGACCAAGAAATACATGCCCGCCTCCCAAGGAGACCTTC 890
 QY 1039 TTATCTTTCGAGATGAGTATAGTCTTCTTACCGCATAAAGCGCTCTCTTCTGAGGTCC 1098
 DB 891 TTATCTTTCGAGATGAGTATAGTCTTCTTACCGCATAAAGCGCTCTCTTCTGAGGTCC 950
 QY 1099 CATAGAAATGAAGAACTGAACGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1158
 DB 951 CATAGAAATGAAGAACTGAACGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1010
 QY 1159 AACCATGGCCATGAAATATCAACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218
 DB 1011 AACCATGGCCATGAAATATCAACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
 QY 1219 TTATTTCAAGA----- 1229
 DB 1071 TTATTTCAAGATGGAACAATAGGAGAAACATCAACATTTGCAATTTGTAGGTCTGATCTT 1130
 QY 1230 ----- 1229
 DB 1131 CTAGAAAGTGAACGCGAGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1190
 QY 1230 -----TGGATTGATGGGAAAGATGGGACTCGTTCATGACCAAGCACTTACGCTT 1272
 DB 1191 TTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTTCATGACCAAGCACTT 1250
 QY 1273 TTAATGAGTCTGGATATATCTCTGAGAGATGAGGCTGTCGACACTTTGGGATTT 1332
 DB 1251 TTAATGAGTCTGGATATATCTCTGAGAGATGAGGCTGTCGACACTTTGGGATTT 1310
 QY 1333 GCACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGACGATTAATGAAGATACAGT 1392
 DB 1311 GCACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGACGATTAATGAAGATACAGT 1370
 QY 1393 CTTGGACTGCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
 DB 1371 CTTGGACTGCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1430
 QY 1453 GGAGAGGGAACATGTGTAAAGTTCGAGGCGGACATGATGCTTCCCTTACATTTGGCAGGA 1512
 DB 1431 GGAGAGGGAACATGTGTAAAGTTCGAGGCGGACATGATGCTTCCCTTACATTTGGCAGGA 1490
 QY 1513 GCGAATGGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
 DB 1491 GCGAATGGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1550
 QY 1573 ACCGCTCAAGCTATCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632
 DB 1551 ACCGCTCAAGCTATCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1610
 QY 1633 GAGAAATTCGAGGAGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1692
 DB 1611 GAGAAATTCGAGGAGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1670
 QY 1693 AAAGCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
 DB 1671 AAAGCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1730
 QY 1753 CGTATTCGAAAGGAAATGTGAGACTTAAATTTATGCTGAGCAGCAGAGGCAAAATTTGTGGG 1812
 DB 1731 CGTATTCGAAAGGAAATGTGAGACTTAAATTTATGCTGAGCAGCAGAGGCAAAATTTGTGGG 1790
 QY 1813 CATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1872
 DB 1791 CATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1850
 QY 1873 ACCCATGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932
 DB 1851 ACCCATGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1910
 QY 1933 GGATATCTCATAGGAGTCCCTCTGCAACCAACCCCAAGCCATCGCATGGAGGAGTCTC 1992

CC atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes,
 CC thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 CC glucose transport defect, glomerulonephritis, hypercalcemia, polycystic
 CC kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 CC congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 CC muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 CC urinary system disorders, osteoporosis, dental disease and infection,
 CC growth and reproductive disorders, hypogonadism, fertility, and/or other
 CC pathologies and disorders, viral, bacterial, or parasitic infections.
 CC The present nucleic acid sequence encodes a NOVX protein of the
 XX invention.

SQ Sequence 2432 BP; 567 A; 655 C; 674 G; 536 T; 0 other;

Query Match 61.6%; Score 1783.6; DB 24; Length 2432;
 Best Local Similarity 81.9%; Pred. No. 0;
 Matches 2374; Conservative 0; Mismatches 34; Indels 490; Gaps 8;

QY 1 CGCTCTCGATGAAGCCCGCGCGCGGATCGCGGCGCTTGGCGGCTGTGGATGCTG 60
 DB 22 CGCTCTCGATGAAGCCCGCGCGCGGATCGCGGCGCTTGGCGGCTGTGGATGCTG 81
 QY 61 TTGGCGCAGGTCCGCGAGCAGGTGAGTCCCGGCGCTCCACACGCGGAAACGCGGG 120
 DB 82 CTGGCGCAGGTCCGCGAGCAGGTGAGTCCCGGCGCTCCACACGCGGAAACGCGGG 141
 QY 121 TCGGACAGCTCGAGCGAGTCCCGCGGCTCTCCCGGCGACCGCGCTCACC 180
 DB 142 TCGGACAGCTCGAGCGAGTCCCGCGGCTCTCTCCCGGCGACCGCGCTCACC 201
 QY 181 GCGATGTCGCGCTGTTTCCGAGGCGACCTGCGTGGCGCATGGGACCGCGGCGAGC 240
 DB 202 GCGATGTCGCGCTGTTTCCGAGGCGACCTGCGTGGCGCATGGGACCGCGGCGAGC 261
 QY 241 GCTTGGAGCCCGAGGCTCCCGCTCTCTCCACCGCGGCGCGCGGCTGGATGGA 300
 DB 262 GCTTGGAGCCCGAGGCTCCCGCTCTCTCCACCGCGGCGCGCGGCTGGATGGA 321
 QY 301 AAAGGCGGGACATGATGAAGCTGGAAACCACTGTTCTCAGCAAACTAACACAGGAACA 360
 DB 322 AAAGGCGGGACATGATGAAGCTGGAAACCACTGTTCTCAGCAAACTAACACAGGAACA 381
 QY 361 GAAACCAAGACTGATGTTCTCAGTC---AATATGACCTGGTCTCTGCCCTACGAGTT 417
 DB 382 GAAACCAAGACTGATGTTCTCAGTCGTGAATATGACCTGGTCTCTGCCCTACGAGTT 441
 QY 418 GACACAGGCGGATTAAGTGTCCCATGAATCATGCACCATCAGCGCGGAGAGACA 477
 DB 442 GACACAGGCGGATTAAGTGTCCCATGAATCATGCACCATCAGCGCGGAGAGACA 501
 QY 478 GTGGCGGTGTCGAGTTGAGTCTTACCTTCGGCTGAAGAGGCCCGACGACCTTC 537
 DB 502 GTGGCGGTGTCGAGTTGAGTCTTTCACCTTCGGCTGAAGAGGCCCGACGACCTTC 561
 QY 538 CACATGATCTGAGGACTTCCAGCAGCTAGTGGCTCTGCTGCTTATTTGTGAGAGCTTG 597
 DB 562 CACATGATCTGAGGACTTCCAGCAGCTAGTGGCTCTGCTGCTTATTTGTGAGAGCTTG 621
 QY 598 GGAAGACAGGCACTAAGTCTGTGCAGACTTTACCGCCAGAGGACTTCTGTTCTATCAA 657
 DB 622 GGAAGACAGGCACTAAGTCTGTGCAGACTTTACCGCCAGAGGACTTCTGTTCTATCAA 681
 QY 658 GGCTCTTTGGATCACACAGAACTCCCATCGATGAGGAGGAGTTCTGTGAGGCTCC 717
 DB 682 GGCTCTTTGGATCACACAGAACTCCCATCGATGAGGAGGAGTTCTGTGAGGCTCC 741
 QY 718 ACTCGACTCTGAGCTCTGCAACAGTCTGAAATGTCCTCCCGGACAGTGTGACCTCCGT 777
 DB 742 ACTCGACTCTGAGCTCTGCAACAGTCTGAAATGTCCTCCCGGACAGTGTGACCTCCGT 801
 QY 778 GCTGCTCTGAGTGTCCGAGCAGACAGACAGGATTCAGAGGCGGCGCATACAGTGGAG 837
 DB 802 GCTGCTCTGAGTGTCCGAGCAGACAGACAGGATTCAGAGGCGGCGCATACAGTGGAG 861

QY 838 CTTACACTCAAGTAGAGCCGACTTATGCAAACTCTACTGTATCCGAGAGATTTCAT 897
 DB 862 CTTACACTCAAGTAGAGCAGGACTTATGCAAACTCTACTGTATCCGAGAGATTTCAT 921
 QY 898 TTCTCTTTCTTTTGTCAATAAAGTCAAAAGATGGGACTCCATGCTCCGAGAGATAGCGT 957
 DB 922 TTCTCTTTCTTTTGTCAATAAAGTCAAAAGATGGGACTCCATGCTCCGAGAGATAGCGT 981
 QY 958 AATGTTTGTATAGATGGGATATGTGAGCTGAGTGTGGTGTCCACATCTCTGCGCACATGCC 1017
 DB 982 AATGTTTGTATAGATGGGATATGTGAG-----ATGCCC 1014
 QY 1018 CAGCTCCCAAGAGAGACCTCTTCATCTTTCAGATGAGTATAAGTCTTGTTCACGGAT 1077
 DB 1015 CAGCTCCCAAGAGAGACCTCTTCATCTTTCAGATGAGTATAAGTCTTGTTCACGGAT 1074
 QY 1078 AAGCGTCTCTTCTGAGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGT 1137
 DB 1075 AAGCGTCTCTTCTGAGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGT 1134
 QY 1138 GTGCAAAAAGATGATGCAAAACCATGGCCATGAATATATCACCACCTACGCTCAGG 1197
 DB 1135 GTGCAAAAAGATGATGCAAAACCATGGCCATGAATATATCACCACCTACGCTCAGG 1194
 QY 1198 ATACTCAACATGATGATCTCTTATTCAAAGATGAGTATGATGGGAAAGATGGGACTCGT 1257
 DB 1195 ATACTCAACATGATGATCTCTTATTCAAAGATGAAACATAGG----- 1238
 QY 1258 CATGACACGCCATCTTACTGACTGCTCTGGATATATGTTCTCTGGAGATGAGCCCTGT 1317
 DB 1239 ----- 1238
 QY 1318 GACACTTTGGGATTGCAACCATAAGTGAATGTGTAGTAAATATCGCAGCTGACGATT 1377
 DB 1239 -----AGGAACATCAACATTGCAATTGT 1262
 QY 1378 AATGAAGATACAGGCTTGGACTGGCTTCACATGTCCTGATGAGTCTGGACACAACTTT 1437
 DB 1263 AGGTCTGATTC----- 1273
 QY 1438 GGCATGATTCATGATGGAGAGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCC 1497
 DB 1274 ----- 1273
 QY 1498 CTTACATTTGGCAGGACGCAATGAGTCTTCTCTGCTACCTGCGAGCGCGAGTATCTA 1557
 DB 1274 ----- 1273
 QY 1558 CACAAATTTTAAGCACCCTCAAGCTATCTGCCTGTGATCAGCCAAAGCCTGTGAAG 1617
 DB 1274 ----- 1273
 QY 1618 GAATACAAGTATCCTGAGAAATGGCCAGGAAATATATGATGCAAAACACACAGCTGCAAG 1677
 DB 1274 ----- 1273
 QY 1678 TGGCAGTTTGGAGAGAAAGCCAAAGCTCTGCACTGTGAGCTTTTAAAGAGACATCTGTAA 1737
 DB 1274 -----TTCTAGAGATGAACAGAGACATCTGTAAA 1302
 QY 1738 GCCCTGTGGTCCCTGATTTGGAGAGAAATGTGAGACTTAAATTTTATGCCAGCAGCAGAA 1797
 DB 1303 GCCCTGTGGTCCCTGATTTGGAGAGAAATGTGAGACTTAAATTTTATGCCAGCAGCAGAA 1362
 QY 1798 GGCACAAATTTGTGGGATGACATGTGCTCCCGGAGGAGCAGTGTGTGAAATATGTTGAT 1857
 DB 1363 GGCACAAATTTGTGGGATGACATGTGCTCCCGGAGGAGCAGTGTGTGAAATATGTTGAT 1422
 QY 1858 GAGGCGCCCAAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1917
 DB 1423 GAGGCGCCCAAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482

QY 1918 AGGACCTCGCGAGGGGAGTATCTCATAGAGTGGCTCTGCACCAACCCCAAGCCATCG 1977
DB 1483 AGGACCTCGCGAGGGGAGTATCTCATAGAGTGGCTCTGCACCAACCCCAAGCCATCG 1542
QY 1978 CATGAGGGAGTCTCTGAGGGCTCCACTCGACTCTGAAGCTCTGCAACAGTCTGAGAA 2037
DB 1543 CATGAGGGAGTCTCTGAGGGCTCCACTCGACTCTGAAGCTCTGCAACAGTCTGAGAA 1602
QY 2038 TGTCCTCGGACAGTGTGCTCCGCTGCTCAGTGTCCGCGACACACACAGCAGA 2097
DB 1503 TGTCCTCGGACAGTGTGCTCCGCTGCTCAGTGTCCGCGACACACACAGCAGA 1662
QY 2098 TTCAGAGGGCGGCACTACAGTGGAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGC 2157
DB 1663 TTCAGAGGGCGGCACTACAGTGGAGCC-----TCAGGACTTATGC 1704
QY 2158 AACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTGTCATAAATAGTCAAA 2217
DB 1705 AACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTGTCATAAATAGTCAAA 1764
QY 2218 GATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGA 2277
DB 1765 GATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAG--- 1821
QY 2278 GTTGGATGTGACAAATGCTCTTGTGATGCTGTGTAAGACGCTGTGTGGGGTGTGTAAC 2337
DB 1822 ---GGATGTGACAAATGCTCTTGTGATGCTGTGTAAGACGCTGTGTGGGGTGTGTAAC 1878
QY 2338 GGAATAACTCAGCTGCAGATTCACAGGGGCTCTACACCAAGCACCAACCAAC 2397
DB 1879 GGAATAACTCAGCTGCAGATTCACAGGGGCTCTACACCAAGCACCAACCAAC 1938
QY 2398 CAGTATTATCAGTGTGACCATTTCTTGTGGAGCCCGAGTATCGCATCTATGAATG 2457
DB 1939 ---TATTATCAGTGTGACCATTTCTTGTGGAGCCCGAGTATCGCATCTATGAATG 1995
QY 2458 AAGCTCTACCTCTACATTTCTGTGGCAATGTCCTCAGAGGCTACTACCTGAATGG 2517
DB 1996 AAGCTCTACCTCTACATTTCTGTGGCAATGTCCTCAGAGGCTACTACCTGAATGG 2055
QY 2518 CACTGGACCTGACTGGCCCGCGGTACAAATTTTCGGGCACTACTTTTCGACTACAGA 2577
DB 2056 CACTGGACCTGACTGGCCCGCGGTACAAATTTTCGGGCACTACTTTTCGACTACAGA 2115
QY 2578 CGTCTCTATGAGCCGAGAACTTAATCGCTACTGGACCAACCAAGAGACACTGAT 2637
DB 2116 CGTCTCTATGAGCCGAGAACTTAATCGCTACTGGACCAACCAAGAGACACTGAT 2175
QY 2638 GTGGAGCTGTGTTTTCAGGAGGAAACCCGGGTGTTCCCTGGGAATACTCCATGCCGCG 2697
DB 2176 GTGGAGCTGTGTTTTCAGGAGGAAACCCGGGTGTTCCCTGGGAATACTCCATGCCGCG 2235
QY 2698 TTGGGACCGAGAGAGCCCTCGCCACCCAGCTACACTTGGGCCCATCTGCGCTCT 2757
DB 2236 TTGGGACCGAGAGAGCCCTCGCCACCCAGCTACACTTGGGCCCATCTGCGCTCT 2295
QY 2758 GAGTGTCTCGTGTCTCGGAGGGGTAGTGTCTTCCAGTGTCTCTGAGGAGCAGA 2817
DB 2296 GAGTGTCTCGTGTCTCGGAGGGGTAGTGTCTTCCAGTGTCTCTGAGGAGCAGA 2355
QY 2818 TGTGAGCTTTCAGCCACTGGTACATTTGCACTGGCTTTCTTGAATCTTAATGAGCAGC 2877
DB 2356 TGTGAGCTTTCAGCCACTGGTACATTTGCACTGGCTTTCTTGAATCTTAATGAGCAG-C 2414
QY 2878 CGGGGCTTCTCTCTGCA 2895
DB 2415 CGGGGCTTCTCTCTGCA 2432

RESULT 11
ABS59325

ID ABS59325 standard; DNA; 2902 BP.

XX

AC ABS59325;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human ADAM-TS 7-like gene #1.
XX
DE Human;
KW Human; NOVA; cardiomyopathy; atherosclerosis; cell signal processing;
KW breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
KW anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
KW neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
KW autoimmune disease; allergy; addiction; asthma; transplantation;
KW graft versus host disease; systemic lupus erythematosus; scleroderma;
KW psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
KW atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
KW thrombocytopenia; bleeding disorder; metabolic disorder; obesity;
KW glucose transport defect; glomerulonephritis; hypercalcaemia;
KW polycystic kidney disease; renal tubular acidosis; skin disorder;
KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
KW haematopoietic disorder; urinary system disorder; osteoporosis; ds;
KW dental disease; dental infection; growth disorder; reproductive disorder;
KW hypogonadism; fertility disorder; viral infection; bacterial infection;
KW parasitic infection; metabolic pathway modulation; gene therapy; gene;
KW zinc metalloprotease; ADAM-TS 7; alpha-2-macroglobulin precursor;
KW ileal sodium/bile acid cotransporter; prohibitin; MT; CIP4; spinesin;
KW macrophage stimulating protein precursor; fatty acid-binding protein;
KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.
XX
OS Homo sapiens.
XX
PN WO200233087-A2.
XX
PD 25-APR-2002.
XX
PF 17-OCT-2001; 2001WO-US32496.
XX
PR 17-OCT-2000; 2000US-241040P.
PR 17-OCT-2000; 2000US-241058P.
PR 17-OCT-2000; 2000US-241063P.
PR 17-OCT-2000; 2000US-241243P.
PR 20-OCT-2000; 2000US-242152P.
PR 23-OCT-2000; 2000US-242482P.
PR 23-OCT-2000; 2000US-242611P.
PR 24-OCT-2000; 2000US-242880P.
PR 24-OCT-2000; 2000US-242881P.
PR 29-DEC-2000; 2000US-259028P.
PR 20-FEB-2001; 2001US-269813P.
PR 25-APR-2001; 2001US-286324P.
PR 29-MAY-2001; 2001US-294108P.
PR 09-JUL-2001; 2001US-303698P.
PR 16-OCT-2001; 2001US-0981151.
XX
(CURA-) CUPAGEN CORP.

Edinger S, Gerlach V, MacDougall JR, Maliyankar UM, Smithson G;
Millet I, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shinkets RA;
Padigaru M, Guo X, Patturajan M, Taupier RJ, Burgess CE;
Zerhusen BD, Kekuda R, Spytek KA, Gangolli EA, Fernandes ER;
Gorman L;

WPI; 2002-590434/63.
P-PSDB; ABG76896.

Cytoplasmic, nuclear, membrane bound and secreted polypeptides and
nucleic acids encoding the polypeptides for diagnosing and treating
e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and
diabetes -

Claim 8; Page 17; 305pp; English.

The present invention relates to new NOVX (NOV1-10) polypeptides. The
molecules of the invention are useful for treating or preventing a

CC NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or
 CC a disorder related to cell signal processing and metabolic pathway
 CC modulation in humans. NOVX polypeptides, nucleic acids and antibodies
 CC are useful for treating or preventing disorders or syndromes including
 CC breast cancer, Alzheimer's disease, epilepsy, Huntington's disease,
 CC anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis,
 CC neurodegeneration, Parkinson's disease, pain, stroke, autoimmune
 CC disease, allergies, addiction, asthma, endometriosis, graft versus host
 CC disease, systemic lupus erythematosus, scleroderma, transplantation,
 CC psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,
 CC atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes, obesity,
 CC thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 CC glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic
 CC kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 CC congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 CC muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 CC urinary system disorders, osteoporosis, dental disease and infection,
 CC growth and reproductive disorders, hypogonadism, fertility, and/or other
 CC pathologies and disorders, viral, bacterial, or parasitic infections.
 CC The present nucleic acid sequence encodes a NOVX protein of the
 XX invention.

SQ Sequence 2902 BP; 699 A; 752 C; 777 G; 674 T; 0 other;

Query Match 58.5%; Score 1695; DB 24; Length 2902;

Best Local Similarity 78.1%; Pred. No. 0;

Matches 2333; Conservative 0; Mismatches 315; Indels 339; Gaps 11;

QY 40 TTGGCGGCGTGTGGATGCTGTGGCGCAGGTGGCGCAGGTGAGTCCCGGCGCTCC 99
 DB 1 TTGGCGGCGTGTGGATGCTGTGGCGCAGGTGGCGCAGGTGAGTCCCGGCGCTCC 60
 QY 100 CACGACGCGGAAACCGCGGTCCGACAGCTGGAGCGAGTCCCGCGGCTCCTCC 159
 DB 61 CACGACGCGGAAACCGCGGTCCGACAGCTGGAGCGAGTCCCGCGGCTCCTCC 120
 QY 160 CGCGGACCGCGCTCTCCCGGATGTCCCGCTCTTTCCGACGCGACCTCCGCGC 219
 DB 121 CGCGGACCGCGCTCTCCCGGATGTCCCGCTCTTTCCGACGCGACCTCCGCGC 180
 QY 220 CATGGGACCGCGCGCGCTGGAGCGCGCTCCCGCTCTCTCCACCGC 279
 DB 181 CATGGGACCGCGCGCGCTGGAGCGCGCTCCCGCTCTCTCTCC 235
 QY 280 GGAGCGCGCGGTGGATGGAAAGCGCGGACATGGATGAAGCTGGAAACCATCGTCT 339
 DB 236 ----- 235
 QY 340 CAGCAACTAACACAGGACGAGAACCAACCAACTGCTGCTCACTCAATATGACCTG 399
 DB 236 -----GACCTG 241
 QY 400 GTCTCTGCTACGAGTTGACCAAGGCGGCGATAGTGTCCCATGAATCATGACCAT 459
 DB 242 GTCTCTGCTACGAGTTGACCAAGGCGGCGATAGTGTCCCATGAATCATGACCAT 301
 QY 460 CAGCGCGGAGGACAGCTGGCGCTGTCCGAGTTGAGTCTCTTCCACCTTCCGCTGAAA 519
 DB 302 CAGCGCGGAGGACAGCTGGCGCTGTCCGAGTTGAGTCTCTTCCAGGATATGC 361
 QY 520 GGCGCCAGGACGACCTTCCACATGGATCTGAGGACTTCCAGCAGCGCTAGTGGCTCTGCG 579
 DB 362 AGAGCCAGAGAGCTCAGACT-----GTGTGTGGAGGCGCTTCCATGCTAATTTTCAGC 416
 QY 580 TTTATTGTGACAGTTGGGAAAGACAGGACCTAAGTCTGTGACAGACTTTACCGCCAGAG 639
 DB 417 CGGGGTTTTGAACCTTCCAAATGTTCCGCTCTCACTGGAGGAAACAGCATGCTTCCAGA 476
 QY 640 GACTTCTGTTTATCAAGCTCTTTGCGATCAACAGAAACTGCGCCATCGGATGAGGG 699
 DB 477 GAATAATAACAATGCAA-----TGTGTGGAGATCGGCGCTTCAACCCAGAAAG 529
 QY 700 AAGTTCTGTGAGGCGTCCACTGCACTCTGAAGCTCTGCAACAGTCAGAAATGTCCTCCG 759

DB 530 TCTAATGTGTTTCTTTCTTTCTTTTATTTTCCAGTCAGGCATGATGCAACAGAA 589
 QY 760 GACAGTTGTGACTTCGCTGCTCAGTGTGCGGACCAACAGCAGAGAGTTCAGAGG 819
 DB 590 GAGGAGATTACTTCTTAAGGCACTTCCTTACACCTCTCATGGAATCTCGGAGAGCT 649
 QY 820 CGGCACTACAAGTGG--AAGCCTTACACTCAAGTAGAAGCCGACTTATCAAACTCTACT 877
 DB 650 GCCAAGGCGAGCTCGCCATCCACGCTACTGTACAAGAGAGAGTCTCTGGTGAACCTCAAG 709
 QY 878 GTATCGCAGAGGATTTGATTTCTTCTTTTGTCAATAAAGTCAAGATGGGACTC 937
 DB 710 ACATGGGAGC-----TGGCACAATCAACCCCTCCACAGCAGCGA 747
 QY 938 CATGCTCGGAGGATAGCCGTATGTTGTATAGATGGGATATGTGAGCTCAGTCTGCTGT 997
 DB 748 CTTCCGCTGGGACTGCCAACAAGAGAGATTTCTGTGAAGACGCAAGAAAT----- 800
 QY 998 CCACATCTCGGCACATGCCCCAGCCTCCCAAGGAAGACCTTTCATCTTGGCAGATGAGT 1057
 DB 801 -----ACATGCCCGCAGCCTCCCAAGGAAGACCTTTCATCTTGCAGATGAGT 848
 QY 1058 ATAAGTCTTGTACGGCATTAAGCCGCTCTTCTGAGGTCCCATGAAATGAAGAACTGA 1117
 DB 849 ATAAGTCTTGTCTACGGCATTAAGCGCTCTCTTCTGAGGTCCCATGAAATGAAGAACTGA 908
 QY 1118 AGCTGGAGACCTTTGGTGGTGGTGCACAAAGATGATGCACAAACCATGCCCATGAAATA 1177
 DB 909 AGCTGGAGACCTTTGGTGGTGGTGCACAAAGATGATGCACAAACCATGCCCATGAAATA 968
 QY 1178 TCACCACTCTGCTCAGCATCTCAACATGCTATCTGCTTTATTTCAAGA----- 1229
 DB 969 TCACCACTCTGCTCAGCATCTCAACATGCTATCTGCTTTATTTCAAGATGAACAA 1028
 QY 1230 ----- 1229
 DB 1029 TAGGAGGAACATCAACATTCGAATTTGAGTCTGATTTCTTAGAGATGAACAGCCAG 1088
 QY 1230 -----TG 1231
 DB 1089 GACTGGTGAATGTCACACGACACACACCTTAAGTAGTTTCTGCCAGTGGCAGTCTG 1148
 QY 1232 GATTGATGGGAAAGATGGGACTCGTCTATGACACCGCCATCTTACTGACTGCTGGATA 1291
 DB 1149 GATTGATGGGAAAGATGGGACTCGTCTATGACACCGCCATCTTACTGACTGCTGGATA 1208
 QY 1292 TATGTTCTGGNAGATGAGCCCTGTGACACTTTGGGATTTGCAACCCATTAAGTGAATGT 1351
 DB 1209 TATGTTCTGGNAGATGAGCCCTGTGACACTTTGGGATTTGCAACCCATTAAGTGAATGT 1268
 QY 1352 GTAGTAATATATCGCAGCTGCAGATTAATGAAGATACAGGTCTTGGACTGGCCTTCACCA 1411
 DB 1269 GTAGTAATATATCGCAGCTGCAGATTAATGAAGATACAGGTCTTGGACTGGCCTTCACCA 1328
 QY 1412 TTGCCCATGAGTCTGGACACAACTTTGGCATGATTTGATGATGGAGAGGAACTATGTA 1471
 DB 1329 TTGCCCATGAGTCTGGACACAACTTTGGCATGATTTGATGATGGAGAGGAACTATGTA 1388
 QY 1472 AAAAGTCCGAGGGAACATCATGCTCCCTACATTTGGGAGGAGCAATGAGTCTTCTCT 1531
 DB 1389 AAAAGTCCGAGGGAACATCATGCTCCCTACATTTGGGAGGAGCAATGAGTCTTCTCT 1448
 QY 1532 GGTCAACCTTCAGCGCGCCAGTATCTACAAATTTCTAAGCACCGCTCAAGCTATCTGCC 1591
 DB 1449 GGTCAACCTTCAGCGCGCCAGTATCTACAAATTTCTAAGCACCGCTCAAGCTATCTGCC 1508
 QY 1592 TTGCTGATCAGCCAAAGCCTGTGAAGGAATACAAGTATCTCTGAGAAATTCAGAGAGAT 1651
 DB 1509 TTGCTGATCAGCCAAAGCCTGTGAAGGAATACAAGTATCTCTGAGAAATTCAGAGAGAT 1568
 QY 1652 TATATGATGCAACACAGTGTGAAGTGGAGGAGGAAAGCCAGTCTGCTGATGC 1711

antagonist are useful for treating a disease or condition associated with decreased or overexpression of functional PRTS in a patient. The PRTS protein is also useful as an immunogen for preparing polyclonal or monoclonal antibodies by hybridoma technology. An antibody that binds the PRTS proteins is useful for detection and purification of the proteins and can be used to diagnose a condition or disease associated with expression of PRTS in a subject or in a biological sample. The sequences of the invention are useful for diagnosis, treatment and prevention of gastrointestinal disorders such as gastritis, ulcerative colitis, Reye's syndrome, etc; cardiovascular such as atherosclerosis, hypertension, myocardial infarction, etc; autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), asthma, Grave's disease, etc; cell proliferative disorders such as hepatitis, psoriasis, leukemia etc; developmental disorders such as Cushing's syndrome; epithelial disorders such as dermatitis, scabies, eczema, etc; neurological disorders such as Parkinson's disease, dementia, Alzheimer's disease, Huntington's disease, multiple sclerosis, etc; or reproductive disorders such as infertility, impotence. The present sequence represents the cDNA encoding the human PRTS4 protein of the invention.

XX SQ Sequence 4888 BP; 1256 A; 1261 C; 1276 G; 1095 T; 0 other;

Query Match 56.6%; Score 1638; DB 24; Length 4888;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 1774; Conservative 0; Mismatches 0; Indels 126; Gaps 1;

QY 1010 ACATGCCCGCCGCTCCCAAGGAGAGACCTCTTCATCTTGCAGAGTACGATATAAGTCTTGCT 1069
 DB 803 ACATGCCCGCCGCTCCCAAGGAGAGACCTCTTCATCTTGCAGAGTACGATATAAGTCTTGCT 862

QY 1070 TAGCGCATAGCGCTCTCTTCTAGGTCCTCCATAGAAATAGAACTGAACTGGAGACCT 1129
 DB 863 TAGCGCATAGCGCTCTCTTCTAGGTCCTCCATAGAAATAGAACTGAACTGGAGACCT 922

QY 1130 TGGTGGTGGTCGACAAAGATGATGCAAAACCAATGCCATGAAATATCACCACCTACG 1189
 DB 923 TGGTGGTGGTCGACAAAGATGATGCAAAACCAATGCCATGAAATATCACCACCTACG 982

QY 1190 TGCTACGATACCAATGGTATCTGCTTTATTCAAAGA----- 1229
 DB 983 TGCTACGATACCAATGGTATCTGCTTTATTCAAAGATGGAACATAGGAGGAACA 1042

QY 1230 ----- 1229

DB 1043 TCAACATTGCAATTGTAGTCTGATCTCTTAGAAGATGAACGCCAGGACTGGTGATAA 1102

QY 1230 -----TGGATTGATGGGA 1243
 DB 1103 GTCAACAGCAGACCAACCTTAAGTAGCTTCTGCCAGTGGCAGCTGGATTGATGGGA 1162

QY 1244 AAGATGGGACTCGTCATGACACAGCCATCTTACTGACTGTCTGGATATATGTTCTCTGGA 1303
 DB 1163 AAGATGGGACTCGTCATGACACAGCCATCTTACTGACTGTCTGGATATATGTTCTCTGGA 1222

QY 1304 AAGATGAGCCCTGTGACACTTTGGGATTTGCACCCATAGTGAATCTGTAGTAATATC 1363
 DB 1223 AAGATGAGCCCTGTGACACTTTGGGATTTGCACCCATAGTGAATCTGTAGTAATATC 1282

QY 1364 GCAGCTGCAGATTAATGAAGATACAGTCTTGGACTGGCTTCCACCATTCGCCCATGAGT 1423
 DB 1283 GCAGCTGCAGATTAATGAAGATACAGTCTTGGACTGGCTTCCACCATTCGCCCATGAGT 1342

QY 1424 CTGGACACACTTTGGCATGATTCATGATGAGAGGGAACATGTGTAAAGTCCGAGG 1483
 DB 1343 CTGGACACACTTTGGCATGATTCATGATGAGAGGGAACATGTGTAAAGTCCGAGG 1402

QY 1484 GCAACATCATGTCCCTACATTTGGCAGGACGATGGAGTCTTCTCTGTGTACCCCTGCA 1543
 DB 1403 GCAACATCATGTCCCTACATTTGGCAGGACGATGGAGTCTTCTCTGTGTACCCCTGCA 1462

QY 1544 GCAGCCAGTATCTACAAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGTGATCAGC 1603

DB 1463 GCAGCCAGTATCTACACAAATTTTAAAGCACCCCTCAAGCTATCTGCCTTGTCTCAGC 1522
 QY 1604 CAAAGCTGTGAAGGAATCAACATATCTCTGAGAAATTTGCCAGGAGATATATATGATGCAA 1663
 DB 1523 CAAAGCTGTGAAGGAATCAACATATCTCTGAGAAATTTGCCAGGAGATATATATGATGCAA 1582
 QY 1664 ACACACAGTGCATAGTGGCAGTTCGGAGAGAAACCAAGCTCTGCATCTGCTGAGCTTTAAAA 1723
 DB 1583 ACACACAGTGCATAGTGGCAGTTCGGAGAGAAACCAAGCTCTGCATCTGCTGAGCTTTAAAA 1642
 QY 1724 AGACATCTGTAAAGCCCTGTGTGGTGCATCTGTTTGGAGAAATGTGAGACTTAATTTA 1783
 DB 1643 AGACATCTGTAAAGCCCTGTGTGGTGCATCTGTTTGGAGAAATGTGAGACTTAATTTA 1702
 QY 1784 TGCAGCAGCAGAAAGCACAATTTGTGGCATCAGATGTGTGCCGGGAGGACAGTGTG 1843
 DB 1703 TGCAGCAGCAGAAAGCACAATTTGTGGCATCAGATGTGTGCCGGGAGGACAGTGTG 1762
 QY 1844 TGAATATGTGTATGAAGGCCCAAGCCCAACCCATGCGCCACTGGTCTGGACTTGTCTT 1903
 DB 1763 TGAATATGTGTATGAAGGCCCAAGCCCAACCCATGCGCCACTGGTCTGGACTTGTCTT 1822
 QY 1904 GTTCCCATCTCTCAGAGCCTTGGGAGGGGAGTATCTATAGGAGTGCCTCTGCACCA 1963
 DB 1823 GTTCCCATCTCTCAGAGCCTTGGGAGGGGAGTATCTATAGGAGTGCCTCTGCACCA 1882
 QY 1964 ACCCAAGCCATCGCATGGAGGAGTTCCTGTGAGGGCTCCACTCGCAGCTCTGAAGCTCT 2023
 DB 1883 ACCCAAGCCATCGCATGGAGGAGTTCCTGTGAGGGCTCCACTCGCAGCTCTGAAGCTCT 1942
 QY 2024 GCAACAGTGCAGAAATGTCCCGGGACAGTGTGTGACTTCCGTGTCTCAGTGTGCCGAGC 2083
 DB 1943 GCAACAGTGCAGAAATGTCCCGGGACAGTGTGTGACTTCCGTGTCTCAGTGTGCCGAGC 2002
 QY 2084 ACAACAGCAGACGATTCAGAGGGGCGCCTACAGTGGAGCTTACACTCAAGTATGAAG 2143
 DB 2003 ACAACAGCAGACGATTCAGAGGGGCGCCTACAGTGGAGCTTACACTCAAGTATGAAG 2062
 QY 2144 ATCAGAGCTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTTGT 2203
 DB 2063 ATCAGAGCTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTTGT 2122
 QY 2204 CAAATAAGTCAAAGATGGGACTTCCATGCTCGGAGGATGACCCGTAAATTTGTATAGATG 2263
 DB 2123 CAAATAAGTCAAAGATGGGACTTCCATGCTCGGAGGATGACCCGTAAATTTGTATAGATG 2182
 QY 2264 GGTATGTGAGAGGATTTGATGTGCAATGTCTTGGATCTGATCTGTGTGAGAGCTCT 2323
 DB 2183 GGTATGTGAGAGGATTTGATGTGCAATGTCTTGGATCTGATCTGTGTGAGAGCTCT 2242
 QY 2324 GTGGGTGTGTAAACGGGAATAACTCAGCCTGCACAGATTCACAGGGGTCTCTACACCAAGC 2383
 DB 2243 GTGGGTGTGTAAACGGGAATAACTCAGCCTGCACAGATTCACAGGGGTCTCTACACCAAGC 2302
 QY 2384 ACCACACACCAACCGATTTATCATGTGTGTCACCATTCCTTCTGGAGCCCGAGTATCC 2443
 DB 2303 ACCACACACCAACCGATTTATCATGTGTGTCACCATTCCTTCTGGAGCCCGAGTATCC 2362
 QY 2444 GCATCTATCAAAATGAACGCTCTCTACTCTCTACATTTCTGTGGCAATGCCCTCAGAGT 2503
 DB 2363 GCATCTATGAATGAACGCTCTCTACTCTCTACATTTCTGTGGCAATGCCCTCAGAGT 2422
 QY 2504 ACTACTGAATGGGCACTGGACCGTGGACTGGCCCGCGCGGTACAAAATTTTCGGGCACCTA 2563
 DB 2423 ACTACTGAATGGGCACTGGACCGTGGACTGGCCCGCGCGGTACAAAATTTTCGGGCACCTA 2482
 QY 2564 CTTTGCAGTACAGAGCTCTTATATAGCCCGGAGAACTTAATTCCTACTAGGACCAACCA 2623
 DB 2483 CTTTGCAGTACAGAGCTCTTATATAGCCCGGAGAACTTAATTCCTACTAGGACCAACCA 2542
 QY 2624 ACAGACACTGATTTGTGAGGCTGTCTTTCAGGGAGGAAACCCGGGTGTGCTCGGAT 2683
 DB 2543 ACAGACACTGATTTGTGAGGCTGTCTTTCAGGGAGGAAACCCGGGTGTGCTCGGAT 2602

are applicable in diagnosis and screening compounds for drug compositions in treating diseases due to neovascularisation, diabetic omentopathy, chronic rheumatoid arthritis, angiodysplasia, refractory skin and gastric ulcers and post-operative healing failure, including gene therapy.

CC The gene encoding such polypeptide has conserved repolysyn-type ZN- metallopeptidase domain, disintegrin-like domain and TSR1 (thrombospondin type 1) domain. Its encoded protein is characterised by high expression in ovaries, changes in expression dose depending on the sexual cycle, a decrease in tumour cell and location of the gene on the 5p-syndrome deletion site on chromosome 5p15.2-15.3 (associated with Cri-du-chat syndrome). The present sequence is the expression Plasmid pFastBac1-HT-F01256 used to express DNA encoding an ADAMTS protein.

XX
SQ Sequence 7668 BP; 1904 A; 1934 C; 1922 G; 1908 T; 0 other;

Query Match 54.5%; Score 1578.4; DB 24; Length 7668;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 1; Indels 126; Gaps 1;

QY 1068 CTTACGGCATAAGCGCTCTCTCTGAGGTCCCATAGAAATGAAGAACTGAACCTGGAGAC 1127
DB 4106 CTTACGGCATAAGCGCTCTCTCTGAGGTCCCATAGAAATGAAGAACTGAACCTGGAGAC 4165

QY 1128 CTTGGTGGTGGCGCAAAAGATGATGCAAAACCATGGCCCATGAATATCACCACCTA 1187
DB 4166 CTTGGTGGTGGTGGCGCAAAAGATGATGCAAAACCATGGCCCATGAATATCACCACCTA 4225

QY 1188 CGTGCTCAGCATACTCAACATGATGATCTCTCTTATTTCAAGA- 1229
DB 4226 CGTGCTCAGCATACTCAACATGATGATCTCTCTTATTTCAAGAATGGMAAATAGGAGGAAA 4285

QY 1230 ----- 1229
DB 4286 CATCAACATTGCAATTGTAGTCTGATTTCTTGAAGATGAACAGCAGGACTGGTGAT 4345

QY 1230 -----TGGATTGATGGG 1241
DB 4346 AAGTCACCGCAGACCAACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGG 4405

QY 1242 GAAAGATGGAGCTCGTCATGACCAACGCCATCTTACTAGCTGGTCTGGATATATGTCCTG 1301
DB 4406 GAAAGATGGAGCTCGTCATGACCAACGCCATCTTACTAGCTGGTCTGGATATATGTCCTG 4465

QY 1302 GAAGAAATGAGCCCTGTGACACTTTGGGATTTGCCACCCATGAAGTGGAAATGTAGTAAATA 1361
DB 4466 GAAGAAATGAGCCCTGTGACACTTTGGGATTTGCCACCCATGAAGTGGAAATGTAGTAAATA 4525

QY 1362 TCGCAGCTGACGATTAATGAGATACAGGTCTTGGACTGGCTTGCACCATTTGCCCATGA 1421
DB 4526 TCGCAGCTGACGATTAATGAGATACAGGTCTTGGACTGGCTTGCACCATTTGCCCATGA 4585

QY 1422 GTCTGGACACAACTTTGGCATGATTCATGATGGAGAGGGAACATGTGTAAAGATCCGA 1481
DB 4586 GTCTGGACACAACTTTGGCATGATTCATGATGGAGAGGGAACATGTGTAAAGATCCGA 4645

QY 1482 GGGCAACATCTGTGCCCTACATTTGCGAGAGCGCAATGGAGTCTTCTCTGTGTCACCTG 1541
DB 4646 GGGCAACATCTGTGCCCTACATTTGCGAGAGCGCAATGGAGTCTTCTCTGTGTCACCTG 4705

QY 1542 CAGCCGCCAGTATCTACAAATTTCTAAGCACCCTCAAGCTATCTGCTTCTGCTGATCA 1601
DB 4706 CAGCCGCCAGTATCTACAAATTTCTAAGCACCCTCAAGCTATCTGCTTCTGCTGATCA 4765

QY 1602 GCCAAAGCCTGTGAAGGAATACAAATGATCTCTGAGAAATTCGAGGAGAAATATATGATGC 1661
DB 4766 GCCAAAGCCTGTGAAGGAATACAAATGATCTCTGAGAAATTCGAGGAGAAATATATGATGC 4825

QY 1662 AAACACAGTGAAGTGGAGTTCGGAGAGAGCCAGACTCTGATCTGCTGACTTTAA 1721
DB 4826 AAACACAGTGAAGTGGAGTTCGGAGAGAGCCAGACTCTGATCTGCTGACTTTAA 4885

QY 1722 AAAGGACATCTGTAAAGCCCTGTGGTGGCCATCTGATTTGGAGGAAATGTGAGACTAAATT 1781

ABK90322 standard; DNA; 1668 bp.	Matches 1405; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
ABK90322;	
21-OCT-2002 (first entry)	
DNA encoding human ADAM-TS-like zinc metalloprotease.	
ADAM TS-like zinc metalloprotease; Gene; human; ds; TNF-antagonist;	
CD8-antagonist; COPD; congestive heart failure; myocardial infarction;	
ischaemic heart disease; arrhythmia; hypertension; vascular disease;	
cerebrovascular disorder; Parkinson's disease; motor neuron disease;	
dementia; multiple sclerosis; stroke; Alzheimer's disease; dementia;	
Huntington's disease; Creutzfeldt-Jakob disease; dementia; schizophrenia;	
psychosis; attention deficit disorder; chromosome 5.	
Homo sapiens.	
Key	Location/Qualifiers
CDS	1..1668
	/*tag= a
	/product= "ADAM-TS-like zinc metalloprotease"
	/partial
	/note= "No start or stop codons shown"
WO200257461-A2.	
25-JUL-2002.	
09-JAN-2002; 2002WO-EP00129.	
18-JAN-2001; 2001US-262034P.	
(PARB) BAYER AG.	
Zhu Z;	
WPI; 2002-575457/61.	
P-PSDB; ABG31503.	
New isolated polynucleotide encoding an ADAM-TS-like zinc	
metalloprotease polypeptide, useful for preventing, ameliorating or	
treating dysfunctions such as COPD, CNS and cardiovascular disorders	
Claim 1; Fig 1; 119pp; English.	
This invention relates to a novel isolated polynucleotide encoding an	
ADAM-TS-like zinc metalloprotease polypeptide. The protein of the	
invention may have cardiant, asotropic, antiarrhythmic, hypotensive,	
antiparkinsonian, nootropic, neuroprotective, cerebroprotective,	
anticonvulsant, anti-HIV, neuroleptic and tranquiliser activity.	
The DNA sequence of the invention may be used in gene therapy as a	
TNF-antagonist or as a CD8-antagonist. The invention also comprises	
an expression vector containing the ADAM-TS-like zinc metalloprotease	
and reagents that modulate and/or regulate the activity of an ADAM-TS-	
like zinc metalloprotease. These reagents can be used in the manufacture	
of a medicament for disorders such as COPD, cardiovascular disease	
(congestive heart failure, myocardial infarction, ischaemic heart	
disease, atrial and ventricular arrhythmias, hypertension and peripheral	
vascular diseases) and cerebrovascular disorders (Parkinson's disease,	
corticobasal degeneration, motor neuron disease, Huntington's disease,	
sclerosis, stroke, Alzheimer's disease, dementia, multiple	
Creutzfeldt-Jakob dementia, schizophrenia, psychosis and attention can	
deficit disorders). Pharmaceutical compositions of the invention can	
also be used for screening and diagnostic assays. The present	
sequence represents the DNA encoding the human ADAM-TS-like zinc	
metalloprotease of the invention, this gene is located on human	
chromosome 5.	
Sequence 1668 BP; 462 A; 392 C; 426 G; 388 T; 0 other;	
Query Match 47.7%; Score 1381.8; DB 24; Length 1668;	
Best Local Similarity 98.9%; Pred. No. 0;	

QY	2310	TGTTGAACACGTCTGTGGGTGTGTAACGGGATAA	CTCAGCTGCACGATTTCACAGGG	2369
Db	1323	TGTTGAACACGTCTGTGGGTGTGTAACGGGATAA	CTCAGCTGCACGATTTCACAGGG	1382
QY	2370	TCTCTACACCAAGCACACACCAACCAAGTATTAT	CACATGGTCAACCATTCCTCTGG	2429
Db	1383	TCTCTACACCAAGCACACACCAACCAAGTATTAT	CACATGGTCAACCATTCCTCTGG	1442
QY	2430	AGCCCGGAGTATCCGATCTATGAATGAACGCTCT	TACCTCCTACATTTCTGTGGCAA	2489
Db	1443	AGCCCGGAGTATCCGATCTATGAATGAACGCTCT	TACCTCCTACATTTCTGTGGCAA	1502
QY	2490	TGCCCTCAGAAGGTACTACTGAAATGGGCACCTG	GACCGTGGACTGGCCCGCGGTACAA	2549
Db	1503	TGCCCTCAGAAGGTACTACTGAAATGGGCACCTG	GACCGTGGACTGGCCCGCGGTACAA	1562
QY	2550	ATTTTCGGGCACTACTTTTCGACTACAGACGGTCT	TATATAGACCGCGAGAACTTAATGC	2609
Db	1563	ATTTTCGGGCACTACTTTTCGACTACAGACGGTCT	TATATAGACCGCGAGAACTTAATGC	1622
QY	2610	TACTGGACCAACCAACGACACACTGATTGTGGAGT	GTCTGT 2650	
Db	1623	TACTGGACCAACCAACGACACACTGATTGTGGAGT	TAAGT 1663	

Search completed: October 28, 2003, 19:02:47
Job time : 762 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2003, 18:20:48 ; Search time 5371 Seconds
(without alignments)
13100.253 Million cell updates/sec

Title: US-09-981-151a-7
Perfect score: 2895
Sequence: 1 cgcctcgtgatgaagcccg.....ccgggggttcctccgtgcc 2895

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_esti.*
- 9: gb_esti.*
- 10: gb_est2.*
- 11: gb_est3.*
- 12: gb_est4.*
- 13: gb_est5.*
- 14: gb_est6.*
- 15: em_estfum.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_pbg.*
- 27: em_gss_vf1.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1103.8	38.1	5104	11 AK031314	AK031314 Mus muscu
2	713.2	24.6	4335	11 AK053851	AK053851 Mus muscu
3	460	15.9	4593	11 AK036137	AK036137 Mus muscu
4	456.4	15.8	625	12 BM850160	BM850160 K-BST0130

5	454.4	15.7	623	12	BM845044
6	427	14.7	1092	12	BM808410
7	401.2	13.9	962	13	BM458327
8	368.4	12.7	537	12	BM845406
9	339.2	11.7	698	13	BQ444190
10	338.2	11.7	1122	12	BM906555
11	329.4	11.4	332	10	BF933693
12	308.4	10.7	732	14	CA512127
13	259.8	9.0	430	12	BM844919
14	237	8.2	657	13	BM287251
15	221.6	7.7	3399	11	AK035797
16	200.4	6.9	735	12	BM944992
17	188.6	6.5	915	12	BI686750
18	183.4	6.3	3249	11	AK048600
19	182.6	6.3	3231	11	AK048612
20	168.6	5.8	728	14	CD355731
21	166.2	5.7	694	12	BQ38710
22	130.8	4.5	841	13	EX451412
23	130.8	4.5	3525	11	AK048211
24	130.4	4.5	4489	11	AK028459
25	128.6	4.4	698	12	EG919675
26	128.6	4.4	4266	11	AK086767
27	128.6	4.4	4378	11	AK053198
28	127.8	4.4	635	10	BB620675
29	124	4.3	692	28	AZ995669
30	120.8	4.2	212	10	BB590994
31	117.4	4.1	629	14	BY718254
32	114.6	4.0	328	11	AK078334
33	113	3.9	3981	11	BC042213
34	112.4	3.9	1799	11	BC030442
35	110	3.8	722	29	AG038263
36	110	3.8	872	13	BU378207
37	109.6	3.8	772	29	CNS04TGT
38	109	3.8	975	12	BI645713
39	108.6	3.8	536	29	B2904744
40	108.2	3.7	7694	11	AK030315
41	108	3.7	1201	13	EX336652
42	106	3.7	2482	11	AK040370
43	104.8	3.6	662	10	BB612240
44	104.4	3.6	708	13	BQ571453
45	102.4	3.5	2499	11	AK082475

ALIGNMENTS

RESULT 1
AK031314
LOCUS

DEFINITION

AK031314 5104 bp mRNA linear HTC 05-DEC-2002
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030405121 product:hypothetical Reprolysin
family propeptide containing protein, full insert sequence.

ACCESSION AK031314.1 GI:26082263

VERSION AK031314.1

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409 (6821), 685-690 (2001)
MEDLINE
21085660
PUBLISHED
11217851
REFERENCE

AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
REFERENCE
6 (bases 1 to 4593)

AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Murakatsu, M. and Hayashizaki, Y.

TITLE
Direct Submission
JOURNAL
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES
source
1. .4593
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9630038L21"
/db_xref="taxon.10090"
/clone="9630038L21"
/tissue="carebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
misc_feature
1. .4593
/note="hypothetical Metalloprotease (ADAM type)/reprolysin (M12B) family containing protein (InterPro|IPR001590, evidence: InterPro)"

BASE COUNT
1232 a 1145 c 1118 g 1098 t
ORIGIN
Query Match 15.9%; Score 460; DB 11; Length 4593;
Best Local Similarity 59.9%; pred. No. 6e-117;
Matches 928; Conservative 0; Mismatches 460; Indels 162; Gaps 3;

Qy 1238 TGGGAAAGATGGGACTCGTCATGACACCGCCATCTTACTGACTGTCTGGATATATGTT 1297
|||||
Db 728 TGGGAAAGATGGGACAGACACGACCGCCATCTCTCTCAGGATTTGACATTTGTT 787

Qy 1298 CCTGGAGATGAGCCCTGTGACACTTTGGATTTTGCACCCCATTAAGTGAATGTGTAGTA 1357
|||||
Db 788 CCTGGAGATGAGCCCTGTGACACTTTGGATTTTGCATATAGTGGCATGTGCACTA 847
|||||
Qy 1358 AATATCGAGCTCCACGATTAATGAAGATACAGGTCTTGGAGTGGCTTACCATTTGCC 1417
|||||
Db 848 AGTACCGAGCTGTACCATCAATGAAGACACAGGACTTGGCTGGCTTACCATTTGCC 907
|||||
Qy 1418 ATGAGTCTGACACACACTTTGGCATGATTCATGATGGAGGAGACATGCTGTAAGAAT 1477
|||||
Db 908 ATGAGTCTGACACACACTTTGGCATGATTCATGATGGAGGAGACATGCTGTAAGAAT 967
|||||
Qy 1478 CCGAGGCGCAACATCATGTCTCCCTACATTTGGCAGGACGCAATGGAGTCTTCTCTGGT 1537
|||||
Db 968 CAGAGGCGCAACATCATGTCTCCCTACATTTGGCAGGACGCAATGGAGTCTTCTCTGGT 1027
|||||
Qy 1538 CTTGAGCGCCGAGTATCTACACAAATTTCTAAGCACCCCTCAAGCTATCTCCCTTGTG 1597
|||||
Db 1028 CTTGAGCGCCGAGTATCTACACAAATTTCTAAGCAAAATTCCT----- 1058
|||||
Qy 1598 ATCAGCCAAAGCCTGTGAAGGAATACAGTATCTCTGAGAAATTTGCCAGGAGAAATAT 1657
|||||
Db 1059 ----- 1058
|||||
Qy 1558 ATGCAAAACACACAGTGCAGTGGCAGTTGGGAGAGAAAGCCTCTGCATGTGGACT 1717
|||||
Db 1059 ----- 1058
|||||
Qy 1718 TTAAGAGGACATCTGTAAAGCCCTGTGGTGCATCTGATTTGGAGGAAATGTGAGACTA 1777
|||||
Db 1059 -----CAGGACATTTGCAGTCACTCTGTGTCACCGCGTGGCCACAGATGTGAACCA 1113
|||||
Qy 1778 AATTATGCCAGCAGCAGAGGACAAATTTGTGGGCATGACATGTGGTCCCGGGAGGAC 1837
|||||
Db 1114 AGTTATGCTGTGACAGAGGAGGAGCCCTGTGTGGCTTGA-----TGGTGTGGGCAAGGCC 1168
|||||
Qy 1838 AGTGTGTGAATATGTTGTAAGGCGCCCAAGCCCACTAGGCGCCCGCCCATCTGCTGGT 1897
|||||
Db 1169 AGTGTGTGAAGTCTGGAGAACTAGGCGCCCGCCCATCTGCTGGCAGTGGTGTGCTGCT 1228
|||||
Qy 1898 CTTCTTGGTCCCATGCTCCAGACCTCGCGAGGGAGGATCTCTCATAGAGTCCGCTCT 1957
|||||
Db 1229 CGAAGTGGTCAGAAATGTTCTCGCACTTGGCGCGGAGGAGTCAAGTTCCAGGAGAGAC 1288
|||||
Qy 1958 GCACCAACCCCAAGCCATGCGATGAGGAGGAGTCTGTGAGGGCTCCACTCCGACCTCA 2017
|||||
Db 1289 GTAGTAACCCCAAGCCCTCAGTATGTTGGCAGTACTGTCCAGATCTAGCCGATCTACA 1348
|||||
Qy 2018 AGCTCTGCAACAGTCAAGAAATGTCCTCGGAGCAGTGTGACTTCCGTTGCTGTCAAGTGT 2077
|||||
Db 1349 AGCTGTGCAACATTAACCCCTTGTCTGAGAAATAGTGTGGATTTCCGTGCCCAACAGTGT 1408
|||||
Qy 2078 CCGAGCACAACACACAGCAGATTCAGAGGCGGCGCACTACAAGTGGAGGCGCTTACACTCA 2137
|||||
Db 1409 CAGAGTATAACAACAAGCCCTTCCGTGGATGTTGTACCGTTGGAAACCCCTACACAAAG 1468
|||||
Qy 2138 TAGAGATCAGGACTTATGCAAACTCTCTCTGTATGCGAGAGGATTTGATTTCTTTT 2197
|||||
Db 1469 TTGAGAGGAGAGATCGATGTAAACTCTACTGCAAGCGGAGAACTTTGAGTTTCTTTT 1528
|||||
Qy 2198 CTTTGTCAATAAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGATATGTTGTA 2257
|||||
Db 1529 CCATGTCTGCAAGGTTGAAGATGGGAGCGCCCTTGTTCCTCCACACAGAAATGACGCTGCA 1588
|||||
Qy 2258 TAGATGGATATGTCAGAGAGTGGATGTGACAAATGTCCTTGGATCTGTGCTGTGTAAG 2317
|||||
Db 1589 TCGATGGATTTGTGAATCGTGGGATGTGACCATGAGCTTGGCTCTAAGGGCGTTCTG 1648
|||||
Qy 2318 AGCTCTGTGGGGTGTGTAACGGAATAACTCAAGCTGACGATTCACAGGGGGTCTCTACA 2377
|||||
Db 1649 ATGCATGTGTGTGTTGCAAGGCGGATAACTCAACTTGCAGGTTTATTAAGGCGCTGTACC 1708
|||||

Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 97 row: E column: 10
High quality sequence stop: 623.

FEATURES
Source

1. 623
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SN216-97-E10"
/sex="F"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/cell_line="SNU-216"
/lab_host="Top10P"
/clone_lib="S12SN216"
/note="Organ: Stomach; Vector: pcNS; Site: 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dr-tailed vector. The dr-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okazaki-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."
181 a 140 c 150 g 152 t

BASE COUNT
ORIGIN
Query Match 15.7%; Score 454.4; DB 12; Length 623;
Best Local Similarity 99.8%; Pred. No. 9.7e-116;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1230 TGGATTGATGGGAAAGATGGGACTGCTCATGACACGCCATCTTACTGACTGCTGGA 1289
Db 168 TGGATTGATGGGAAAGATGGGACTGCTCATGACACGCCATCTTACTGACTGCTGGA 227
QY 1290 TATATGTTCTTGGAGATGAGCCCTGTGACACTTGGGATTTGCACCCATAGTGAAT 1349
Db 228 TATATGTTCTTGGAGATGAGCCCTGTGACACTTGGGATTTGCACCCATAGTGAAT 287
QY 1350 GTGTAGTAAATATCGACGTCGACGATTAATGAAGATACAGGCTTGGACTGGCCTTCAC 1409
Db 288 GTGTAGTAAATATCGACGTCGACGATTAATGAAGATACAGGCTTGGACTGGCCTTCAC 347
QY 1410 CATTGCCATGATGTCGACACACTTGGCATGATTCATGATGAGAGGGACATGTC 1469
Db 348 CATTGCCATGATGTCGACACACTTGGCATGATTCATGATGAGAGGGACATGTC 407
QY 1470 TAAAGATCCGAGGCAACATCATGTCCTCCCTACATTTGGCAGGACCAATGGATCTTCTC 1529
Db 408 CAAAGATCCGAGGCAACATCATGTCCTCCCTACATTTGGCAGGACCAATGGATCTTCTC 467
QY 1530 CTGGTCACTGTCAGCCCGCAGTATCTACACAAATTTCTAAGACCCGCTCAGCTATCTG 1589
Db 468 CTGGTCACTGTCAGCCCGCAGTATCTACACAAATTTCTAAGACCCGCTCAGCTATCTG 527
QY 1590 CTTGCTGATCAGCCAAAGCCCTGTGAAGCAATCAAGTATCTCTGAGAAATTCGCAGGAGA 1649
Db 528 CTTGCTGATCAGCCAAAGCCCTGTGAAGCAATCAAGTATCTCTGAGAAATTCGCAGGAGA 587
QY 1650 ATTATATGATGCAAAACACAGATGCAAGTGGCAGTT 1685
Db 588 ATTATATGATGCAAAACACAGATGCAAGTGGCAGTT 623

RESULT 6
BM808410

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

Location/Qualifiers

1. 1092

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5591626"

/lab_host="DH10B"

/clone_lib="NIH MGC 125"

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool of three ovaries, from females ranging in age from 38 to 49 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."

250 a 333 c 295 g 214 t

BASE COUNT

ORIGIN

Query Match 14.7%; Score 427; DB 12; Length 1092;

Best Local Similarity 100.0%; Pred. No. 5.9e-108;

Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2357 CGATTTCACAGGGGCTCTTACACCAAGCAGCACCACACCCACCATATATACATGTC 2416

Db 19 CGATTTCACAGGGGCTCTTACACCAAGCAGCACCACACCCACCATATATACATGTC 78

QY 2417 CCATTCTCTTGGAGCCCGGAGTATCCGATCTATGAATGAAGCTCTCTACCTCTTACA 2476

Db 79 CCATTCTCTTGGAGCCCGGAGTATCCGATCTATGAATGAAGCTCTCTACCTCTTACA 138

QY 2477 TTTCTGTGGCAATGTCCTCAGAGGCTACTACCTGAATGGGCACTGACCGTGGACTGGC 2536

Db 139 TTTCTGTGGCAATGTCCTCAGAGGCTACTACCTGAATGGGCACTGACCGTGGACTGGC 198

QY 2537 CCGGCGGCTACAAATTTTCGGGCACTACTTTCGAGTACAGACGGTCTCTAATGAGCCCG 2596

Db 199 CCGGCGGCTACAAATTTTCGGGCACTACTTTCGAGTACAGACGGTCTCTAATGAGCCCG 258

QY 2597 AGAATCTAATCCTACTCGACCAACCAACGAGACACTGATTGTGGAGACTGCTGTTTCAGG 2656

Db 259 AGAATCTAATCCTACTCGACCAACCAACGAGACACTGATTGTGGAGACTGCTGTTTCAGG 318

QY 2657 GAAGGAACCCGGGTGTCCTGGGAATACCTCCATCGCTTCGCGGACCGAGAGCAGC 2716

BM808410 1092 bp mRNA linear EST 05-MAR-2002
AGENCOURT 6604069 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5591626
5', mRNA sequence.

BM808410
EST.
BM808410.1 GI:19125233

EST.
Homo sapiens (human)

Homosapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1092)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12367 row: n column: 11

High quality sequence stop: 728.

Location/Qualifiers

1. 1092

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5591626"

/lab_host="DH10B"

/clone_lib="NIH MGC 125"

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool of three ovaries, from females ranging in age from 38 to 49 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."

250 a 333 c 295 g 214 t

BASE COUNT

ORIGIN

Query Match 14.7%; Score 427; DB 12; Length 1092;

Best Local Similarity 100.0%; Pred. No. 5.9e-108;

Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2357 CGATTTCACAGGGGCTCTTACACCAAGCAGCACCACACCCACCATATATACATGTC 2416

Db 19 CGATTTCACAGGGGCTCTTACACCAAGCAGCACCACACCCACCATATATACATGTC 78

QY 2417 CCATTCTCTTGGAGCCCGGAGTATCCGATCTATGAATGAAGCTCTCTACCTCTTACA 2476

Db 79 CCATTCTCTTGGAGCCCGGAGTATCCGATCTATGAATGAAGCTCTCTACCTCTTACA 138

QY 2477 TTTCTGTGGCAATGTCCTCAGAGGCTACTACCTGAATGGGCACTGACCGTGGACTGGC 2536

Db 139 TTTCTGTGGCAATGTCCTCAGAGGCTACTACCTGAATGGGCACTGACCGTGGACTGGC 198

QY 2537 CCGGCGGCTACAAATTTTCGGGCACTACTTTCGAGTACAGACGGTCTCTAATGAGCCCG 2596

Db 199 CCGGCGGCTACAAATTTTCGGGCACTACTTTCGAGTACAGACGGTCTCTAATGAGCCCG 258

QY 2597 AGAATCTAATCCTACTCGACCAACCAACGAGACACTGATTGTGGAGACTGCTGTTTCAGG 2656

Db 259 AGAATCTAATCCTACTCGACCAACCAACGAGACACTGATTGTGGAGACTGCTGTTTCAGG 318

QY 2657 GAAGGAACCCGGGTGTCCTGGGAATACCTCCATCGCTTCGCGGACCGAGAGCAGC 2716


```

Db      319 GAAGGAACCCGGGTGTTGCTGGGAATACTCCATGCTCGTGTGGGACCGAGAGCAGC 378
QY      2717 CCCTGCCAGCCAGCCAGCTACACTTGGCCCATCGTGGCTCTGAGTGTCTCCGTGTCCTGGC 2776
Db      379 CCCCTGCCAGCCAGCCAGCTACACTTGGCCCATCGTGGCTCTGAGTGTCTCCGTGTCCTGGC 438
QY      2777 GAGGGGG 2783
Db      439 GAGGGGG 445

RESULT 7
BX458327
LOCUS
DEFINITION BX458327 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE001YN22
5-PRIME, mRNA sequence.
ACCESSION BX458327
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 962)
AUTHORS Li, W. B., Gruber, C., Jesse, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8035.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE001DG11QPI&cluster=8035.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DE001DG11QPI.
FEATURES
    source
        Location/Qualifiers
            1..962
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DE001YN22"
                /tissue_type="PLACENTA"
                /clone_lib="Homo sapiens PLACENTA"
                /notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                with a NotI-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                Library was not normalized."
BASE COUNT 268 a 202 c 249 g 232 t 11 others
ORIGIN
    Query Match 13.9%; Score 401.2; DB 13; Length 962;
    Best Local Similarity 65.9%; Pred. No. 9.7e-101;
    Matches 588; Conservative 5; Mismatches 298; Indels 1; Gaps 1;

QY      1542 CAGCCGCCAGTACTACACAATTTCTAGACCCGCTCAAGCTACTCTGCTGCTGATCA 1601
Db      71 CAGCCGCCAGTACTCTCAAGAAATTTCTCAGCACACCTCAGCGGGGTGTCTAGTGGATGA 130
QY      1602 GCCAAAGCCTGTGAAGAAATACAAAGTATCTCTGAGAATTCGCCAGGAGAAATATATGATGC 1661
Db      131 GCCCAAGCAGCAGGACAGTAAATATCTCGGCAACTACCAAGCAGATTTATGATGC 190
QY      1662 AAACACACAGTGAAGTGGCAGTTCGAGAGAAAGCAAGCTCTCTGATGCTGCACTTTAA 1721
Db      191 TGACACACAGTGTAAATGGCAATTTGGAGCAAAAGCAAGTATATGAGCCTTTGGTTTGT 250
QY      1722 AAAGGACATCTGTAAAGCCCTGTGGTGCCTATCTATTGGAAGAAATGTGAGACTAAAT 1781

```

```

Db      251 GAAGGATATTTCCAAATCACTTTGGTGGCCACCGAGTAGGCCACAGGTGTGAGACCAAGTT 310
QY      1782 TATGCCAGCAGCAGAGGACCAATTTGTGGGCATGACATGTGGTGGCCGGGAGGACAGT 1841
Db      311 TATGCCCGCAGCAGAGGACCGTTTGTGGCTTGAGTATGTGGTGGCAAGGCCAGTG 370
QY      1842 TGTGAAATATGTGATGAAGGCCCCACAGCCACCCATGGCCACCTGGTGGCACTGGTCTTC 1901
Db      371 CGTAAAGTTTGGGGAGCTCGGGCCCGGCCCATCCACGGCCAGTGGTCCGCTGGTCA 430
QY      1902 TTGGTCCCATCTCCAGGACCTCGGAGGGGAGTATCTCATAGGAGTCCCTCTGCAC 1961
Db      431 GTGGTCAGAAATTTCCCGACATGTGGTGGAGGAGTCAAGTTCCAGGAGACACTGCA 490
QY      1962 CAAOCCCAAGCCATCGCATGGAGGAAAGTTCTGTGAGGGCTCCCACTCGCACTCTGAAGCT 2021
Db      491 TAAOCCCAARCCCTCAGTATGTGGCATATTTCTGTCCAGGTTCTAGCCGCTATTTATCAGCT 550
QY      2022 CTGCAACAGTCAGAAATGTCCTCCCGGACAGTGTGACCTTCCTGCTGCTCAAGTGTGCCGA 2081
Db      551 GTGCAATATTAACCTTGCATGAATAATAGCTTGGATTTCCGGCTCAACAGTGTGCAGA 610
QY      2082 GCACACACACACACGATTCAGAGGGCGGCATACAAAGTGGAGGCTTACACTCAAGTAGA 2141
Db      611 ATATAACAGCAAAACCTTTCCGTGGATGTTCTACCAAGTGGAAACCTTATCAAAAGTGA 670
QY      2142 AGTCAAGGACTTATGCMAACTCTCTGATCGCAAGGATTTG-ATTCTCTCTTTCTT 2200
Db      671 AGAGGAAGATCGATGCAAACTGTACTGCAAGCTGGAACCTTTGAATTTTTTTTGGAA 730
QY      2201 TGTCAATATAAGTCAAGATGGGACTTCCATGCTCGGAGGATAGCGGTAACTTTGTATAG 2260
Db      731 TGTCCGGCAAGTGAARAAGGAACTCCCTGCTCCCAACAAACAAAGATGTTGTATG 790
QY      2261 ATGGGATGTGAGAGGATTCGATCTGCAACTGCTCTGATGCTGATGCTCTGATGCTGAGAGC 2320
Db      791 ACGGGGTTTGAAGTCTAGTGGATGTTGATCATGAAGTCTCTTAAAGCTTTCAGATG 850
QY      2321 TCTGTGGGTGTGTAAACGGGAATACTCAGCCTGCAGCATTCACAGGGGTCTCTACACCA 2380
Db      851 CTGTGCGCTTTGCAAGGTGATTAATCACTTCAAGTTGCAAGTTTATAAGGCTGTACCTCA 910
QY      2381 AGCACACACACACACACGATTCATCATGCTGCTCACCATTCTCTGAGC 2432
Db      911 ACCAGCATAAAGCAAAAGAAATATATCCGGTGGTCTCATCTCCAGCTGGCGC 962

RESULT 8
BM845406
LOCUS
DEFINITION K-EST0123712 S12SNU216 Homo sapiens cDNA clone S12SNU216-99-B04 5',
mRNA sequence.
ACCESSION BM845406
VERSION
KEYWORDS BM845406.1 GI:19201805
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 537)
AUTHORS Kim, N. S., Hahn, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Chu, M. Y., Kim, M. R.,
Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and
Kim, Y. S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 99 row: B column: 04

```


High quality sequence stop: 537.
 Location/Qualifiers
 1..537
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S12SN216-99-B04"
 /sex="F"
 /tissue_type="Lymph node"
 /cell_type="Epithelial"
 /cell_line="SNU-216"
 /lab_host="R610F"
 /clone_lib="S12SN216"
 /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dr-selected mRNA by priming with dr-tailed vector. The dr-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli R610F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."
 BASE COUNT 149 a 124 c 129 g 135 t
 ORIGIN
 Query Match 12.7%; Score 368.4; DB 12; Length 537;
 Best Local Similarity 99.7%; Pred. No. 1.2e-91;
 Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1230 TGGATTGATGGGAAAGATGGGACTCGTCATGACACGCCCATCTTACTGACTGTCGGA 1289
 Db 168 TGGATTGATGGGAAAGATGGGACTCGTCATGACACGCCCATCTTACTGACTGTCGGA 227
 QY 1290 TATATGTTCTCGGAAGATGAGCCCTGTGACATTTGGGATTTGACCCCATAGTGGAAAT 1349
 Db 228 TATATGTTCTCGGAAGATGAGCCCTGTGACATTTGGGATTTGACCCCATAGTGGAAAT 287
 QY 1350 GTGTAGTAATATCGACGTCACGATTAATGAAGATACAGGCTTGGACTGGCCCTTCAC 1409
 Db 288 GTGTAGTAATATCGACGTCACGATTAATGAAGATACAGGCTTGGACTGGCCCTTCAC 347
 QY 1410 CATTCGCCATGATCTGGACACACTTTGGCATGATTCATGATGGAGGACATGTC 1469
 Db 348 CATTCGCCATGATCTGGACACACTTTGGCATGATTCATGATGGAGGACATGTC 407
 QY 1470 TAAAGATCCGAGGCAACATCATGTCCCTACATTTGGCAGGACCAATGGAGTCTTCTC 1529
 Db 408 CAAAGATCCGAGGCAACATCATGTCCCTACATTTGGCAGGACCAATGGAGTCTTCTC 467
 QY 1530 CTGCTCACCTTCAGCCCGAGTATCTACCAATTTCTAGCCACCCCTCAAGCTATCTG 1589
 Db 468 CTGCTCACCTTCAGCCCGAGTATCTACCAATTTCTAGCCACCCCTCAAGCTATCTG 527
 QY 1590 CTTTGTCTGAT 1599
 Db 528 CTTTGTCTGAT 537
 RESULT 9
 BQ444190
 LOCUS
 DEFINITION
 UI-M-EX0-bxj-1.10-0-UI.r1 NIH BMAP_EX0 Mus musculus cDNA clone
 IMAGE:5709081 5', mRNA sequence.
 ACCESSION
 BQ444190
 VERSION
 BQ444190.1 GI:21247302
 KEYWORDS
 EST.

SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 698)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..698
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5709081"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP_EX0"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoRI;
 Site 2: NotI; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dr primer containing a NotI site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoRI adaptor, digested with NotI, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the NotI site and the polyA tail
 is GTGCGTGGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES
 source

BASE COUNT
 ORIGIN

Query Match 11.7%; Score 339.2; DB 13; Length 698;
 Best Local Similarity 73.3%; Pred. No. 2.1e-83;
 Matches 434; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
 QY 1238 TGGGAAAGATGGGACTCGTCATGACACGCCCATCTTACTGACTGGTCTGGATATGTT 1297
 Db 107 TGGGAAAGATGGGACGACGACGCCCATCTCTCCACAGGATTTGACATTTGTT 166
 QY 1298 CCTGGAAGAATGAGCCCTGTGACACTTTGGGATTTGACCCCATTAAGTGGAAATGTAGTA 1357
 Db 167 CCTGGAAGAAGCAACCATGTGACACACTAGGATTTGCTCTCTATCAGTGGCATGTGCAGTA 226
 QY 1358 TATATCGCAGCTGCACGATTAATGAGTACAGGCTTGGACTGGCCCTTCACCATTTGCC 1417
 Db 227 AGTACCGAAGCTGTACCATTAATGAAGACACAGGACTTGGCCCTTCACCATTTGCC 286
 QY 1418 ATGAGTCTGGACACAACATTTTGGCATGATTCATGATGAGAGGAGCAACATGTGTAAAGT 1477
 Db 287 ATGAGTCAAGGACACAACATTTTGGCATGATTCATGATGAGAGGCAACATTCCTGAGGAAG 346
 QY 1478 CCGAGGCAACATCATGTCCCTACATGCGACAGACCAATGAGTCTTCTCTGTCAC 1537
 Db 347 CAGAAAGCAACATCATGTGTCACCCCACTGACTGGAAACAATGGGGTGTGTTTCATGTCGT 406

QY 1538 CTTGAGCGCCAGTATCTACACAAATTTCTAAGCACCCTCAAGCTATCTGCTTGTG 1597
 DB 407 CTTGAGCGCCAGTATCTAAGAAATTTCTCAGTACACCGAGGCTGGCTGTCTGGTG 466
 QY 1598 ATCAGCCAAAGCCTGTGAGGAATACAGATCTCTGAGAAATTTGCCAGGGAATATATG 1657
 DB 467 ATGAGCCCAAGCAACCGGACAGTATAAATATCCGGAACAACTCCCGGACAGATTTATG 526
 QY 1658 ATGCAACACACAGTGCAGTGCAGTGTGGAGTGGAGAGAACCAAGCTCTGCACTGCACT 1717
 DB 527 ATGCGAGACATGCAATGTAGTGGCAATTTGAGGCCAAAGCAAGCTGTGAGCCTTGGG 586
 QY 1718 TTAAGAAGGACATCTGAAGCCTGTGGTCCATCTGATTTGAGGAAATGTGAGACTA 1777
 DB 587 TTATGAAGGACATTTGCAAGTACCTCTGTGTCACCGCGTGGGCCACAGATGTGAACCA 646
 QY 1778 AATTATGCCAGCAGCAGAGGCAAAATTTGTGGGCATGACATCTGTGTGGCG 1829
 DB 647 AGTTTCATGCTGCAGCAGAGGGACCGCTTGTGGCTTGAGTATGTGTGTGCG 698

RESULT 10
 BM906555 1122 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT 6621721 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5722286
 DEFINITION 5' mRNA sequence.
 ACCESSION BM906555
 VERSION BM906555.1 GI:19356934
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12708 row: b column: 15
 High quality sequence start: 19
 High quality sequence stop: 633.
 High quality location/Qualifiers

FEATURES

1. 1122
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5722286"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_125"
 /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
 Site 1: ECoRV (destroyed); Site 2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 (ECoRV site is destroyed upon cloning). Average insert
 size 2.1 kb, insert size range 1-3.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036."

BASE COUNT 296 a 306 c 287 g 233 t

ORIGIN

Query Match 11.7%; Score 338.2; DB 12; Length 1122;
 Best Local Similarity 70.6%; Pred. No. 4.8e-83;
 Matches 50; Conservative 0; Mismatches 190; Indels 44; Gaps 8;

QY 428 GCGATTACGTGTCATGAAATCATGCAACCATCAGCGCGGAGAGCAGTGCCTGT 487
 DB 1 GCGATTACGTGTCCTGCTGAAATCATGCAACCATCAGCGCGGAGAGCAGTGCCTGT 60
 QY 498 CCGAGGTTGACTCTCTTCCACCTTCGGCTGGAAGCCCCAGCAGCTTCCACATGATC 547
 DB 61 CCGAGGTTGAGTCTCTTCCACCTTCGGCTGGAAGCCCCAGCAGCTTCCACATGATC 120
 QY 548 TGAGGACTTCCAGCAGCCTAGTGGCTCTCTGCTTTATTTGTGCAGACCTTG--GGAAAGACA 606
 DB 121 TCGGGAATTCAGCAGCCTAGTGGCTCTCTGCTTTATTTGTGCAGACCTTGCGAAGACA 180
 QY 607 GGCATTAAGTCTGTGCAGACTTTACCGCCAGCAGACTCTCTGTTCTATCAAGGCTCTTTG 666
 DB 181 GGCATTAAGTCTGTGCAGACTTTACCGCCAGCAGACTCTCTGTTCTATCAAGGCTCTTTG 240
 QY 667 CGATCACACAGAACTCGCCATCGCATGCGAGGGAAGTTCTGTGAGGGCTTCCACTCGCACT 726
 DB 241 CGATCACACAGAACTCTCTCA-----GTGGCCCTTTTCAACC 276
 QY 727 CTGAAGCTCTGCAACAGTCAGAAATGTCCTCCGGGACAGTGTGACTTCCGTGCTGCTAG 786
 DB 277 TGCCAAAGCTTGTTCAGGCGATGATAGCAACAGAGAGGCGAGATTACTTCTTAAGGCCACTT 336
 QY 787 TGTGCGGAGCACAACAGCAGACGATTCAGAGGCGCGCACTACAAGTGG--AAGCCTTACA 844
 DB 337 CCTTCACACCTCTCATGGAACTCGGAGAGCTCCCAAGCAGCTGCCATCCACGTA 396
 QY 845 CTCAGTGAAGCGGACCTTATGCAAACTCTACTGTATCGCAAGAGAGATTTGATTTCTTCT 904
 DB 397 CTGTACAAGAGATCCACAGAGCCCTCTCTCTGGGCGCAGTGAAGTCTCTGGTCAACCTC- 455
 QY 905 TTCTTTTCAAAATAAGTCAAAGATGGAGCTCCATGCTCGGAGGATAGCGTAACTGTT 964
 DB 456 -----AAGGACATGGAGCTGGCAATCAACCCCTGACAGCAGGACCTTCGC 504
 QY 965 GTATAGATGGGATATGTGAGCTCAGTGTGTGTCCACATCTG--CGCACATGCCCCAGCCT 1023
 DB 505 CTGGGACTGCCAAGAACGAGCATTTTCTGTGGAAGACCAAGAAATACATGCCCCAGCCT 564
 QY 1024 CCCAGGAGAGCTCTTCTCTCTCCAGATGATGATTAAGTCTTCTTACGGCATAAGCC 1083
 DB 565 CCCAGGAGAGCTCTTCTCTCTCCAGATGATGATTAAGTCTTCTTACGGCATAAGCC 624
 QY 1084 TCTCTTCTGAGTCCCATAGAAATCAAGAACTGAACCTGGAGAGCTTGGTGGTGGTGCAC 1143
 DB 625 TCTCTTCTGAGTCCCATAGAAATCAAGAACTGAACCTGGAGAGCTTGGTGGTGGTGCAC 684
 QY 1144 AAAAAAGTATGCAAAA--CCATGCCCATGAATAATATCA--CCACTAGCTGCTCAGATAC 1201
 DB 685 CAAAAGATGATGCAAAAACCCCATGGCCATGAATAATATCAACCCCTACGGGCTCACCATAC 744
 QY 1202 TCACATGGTA--TCTGCTTTTATCAAGATGGATTTGATGGG 1242
 DB 745 TCAACATGGTAATCTGCTTTTATCAAGAGGAGAACATAGG 787

RESULT 11

BF933693/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BF933693 332 bp mRNA linear EST 22-JAN-2001
 IL5-NT0228-211200-358-e08 NT0228 Homo sapiens cDNA, mRNA sequence.
 BF933693
 BF933693.1 GI:12351017
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 332)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

Job time : 5382 secs

Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

Location/Qualifiers

source

1..3399
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9630005B22"
/db_xref="taxon:10090"
/clone="9630005B22"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
264..1757

CDS

/note="unnamed protein product; hypothetical
Metalloprotease (ADAM type)/reprolysin (M12B) family
containing protein (InterPro|IPR001590, evidence:
InterPro)
putative"
/codon_start=1
/protein_id="BAC29190.1"
/db_xref="GI:26330920"
/translation="MECALLCLCALRAAGPMPGAGLAKALQCCFCCASVAV
ALASDGGSGGLNDYVVFVVEVDGSGYISHDIHHRKRSARAGNSLHYRVS
AFQDLHLKPSAILSSHFVQVLKDGASREPEVFCLYOGFIRNDSSSVAVS
TCAGLGLRTDNBFLSLPOLLAQHNYSBPAGHPHVLKRTAEKERVYQDYP
GGRTTPGSPSHPTFPAGSQSEPEYSHRWQKRFHFCGRKKIAPKPAEDNLYRDEY
GGTPRRSAGSNQNLVETLVADAKWVEKHGKDDVTYILVMNVSSLFKDGFI
GSDINIVVSLLEBEPGLLNHHADQSLNSFCQMSALVQNGKXRDHAILLTGF
DSCWNKNEPFDLTGAPISGMCYSKTSCTINEDTGLAFTIAHESGHNFQVHDDGEG
NPKRAEGNIMSPITLGNNGVFSWSSCSQYVLRKFLRTPRCIAFLTG"

BASE COUNT 905 a 838 c 774 g 882 t

ORIGIN

Query Match 7.7%; Score 221.6; DB 11; Length 3399;
Best Local Similarity 76.4%; Pred. No. 3.9e-50;
Matches 272; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY	1238	TGGGAAAGATGGGACTCGTCATGACACGACCATCTTACTGACTGGTCTGGATATATGTT	1297
Db	1390	TGGGAAGATGSCAAGACACGACGACGACCATCTCTCACAGATTGACATTGTT	1449
QY	1298	CCTGGAAGATGAGCCCTGTGACACTTTGGGATTTGACCCCATAGTGGATGTGTAGTA	1357
Db	1450	CCTGGAAGACGAAACCATGTGACACTAGGATTTGCTCTATCATGCGCATGTGCAGTA	1509
QY	1358	AATATCGAGCTCCAGGATTAATGAGATACAGGCTTTGGACTGGCTTCACCATTTGCC	1417
Db	1510	AGTACCGAGCTGTACATCATGACACAGGACTTGGCTGGCTTCACCATTTGCC	1569
QY	1418	ATGAGTCTGACACAACTTTGGCATGATTCATGATGACAGGAAACATGTGTRAAAGT	1477
Db	1570	ATGAGTCAGGGCAACACTTTGGCATGATTCATGATGAGGCAATCCCTGCAGGAAG	1629
QY	1478	CCGAGGCAACATCATGTCCCTACATTTGGCAGGACGCAATGGAGTCTTCTCGGTGAC	1537
Db	1630	CAGAGGCAACATCATGTCCCTACATTTGGCAGGACGCAATGGAGTCTTCTCGGTGAC	1689
QY	1538	CCTGCAAGCCGCAAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCCTT	1593
Db	1690	CCTGCAAGCCGCAAGTATCTAAGAAATTCCTCAGAACTCCTAGATGTATAGCCCTT	1745

Search completed: October 28, 2003, 23:23:11

Db 349 AGCCTAGTGGCTCTGGCTTTATTTGTCAGAGCTTGGGAAAGACAGGACACTAAGCTCTGTG 408
Qy 622 CAGACTTTACCGCAGAGAGACTTCTGTTTCTATCAGAGCTCTTTGCGATCACACAGAAC 681
Db 409 CAGACTTTACCGCAGAGAGACTTCTGTTTCTATCAGAGCTCTTTGCGATCACACAGAAC 468
Qy 682 TCGCCATCGATGAGAGAGAGTCTGTGAGGCTTCACTCGCACTCTGAAGCTCTGCAAC 741
Db 469 TCTCTCA-----GTGGCCCTTTCARCCCTGCAAGGCTTGTCA 504
Qy 742 AGTCAGAAATGTCGCCGACAGTGTGACTTCCGTGCTGTGCTGAGTGTGCGGACACAAAC 801
Db 505 GGCATGATACGAAACAGAGAGGAGATTAATCTCTTAAGGCCACTTCCCTTCACACCTCTCA 564
Qy 802 AGCAGACGATTCAGAGGCGGACACTACAGTGG--AAGCCTTACACTCTAGTGTGAGAGCCG 859
Db 565 TGGAACTCGGACAGCTCCCAAGGAGCTCGCCATCCACGCTACTGTATCAGAGATCC 624
Qy 860 ACTTATGCAAACTCTACTGTATCGCAGAGAGATTTGATTTCTTCTTTTCTTTGTCNAATA 919
Db 625 ACAGAGCCCTCTCTCTGGGCCAGTGAAGTCTCTGTGAGCTC-----AAGG 672
Qy 920 AAGTCAAGATGGACTCCATCTCGAGGATAGCCGTAATGTTGTATAGATGGATAT 979
Db 673 ACATGGAGCTGGCAGATCACTACCTCTGCACGACGACCTTCGCGTGGGCTGCCCAA 732
Qy 980 GTGAGCTCAGTGTGTGTCCACATCTGC--GCATGCCCCAGCCTCCCAAGAGACCTC 1038
Db 733 AAGCAGACTTCTGTGGAAGAGCGCAAGAAATACATGCCCCAGCCTCCCAAGAGAGACCTC 792
Qy 1039 TTCACTTCCGAGATGATTAAGTCTTGTCTACGCGATAGAGCTCTCTCTGAGCTCC 1098
Db 793 TTCACTTCCGAGATGATTAAGTCTTGTCTACGCGATAGAGCTCTCTCTGAGCTCC 852
Qy 1099 CATAGAAATGAAGAACTGAACTGGAGACTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1158
Db 853 CATAGAAATGAAGAACTGAACTGGAGACTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 912
Qy 1159 AACATGCGCATGAAATATCAACCTACCTGCTCAGTACTCACTCACTCACTCACTCACTCACT 1218
Db 913 AACATGCGCATGAAATATCAACCTACCTGCTCAGTACTCACTCACTCACTCACTCACTCACT 972
Qy 1219 TTATTCAAAAGA----- 1229
Db 973 TTATTCAAAAGATGGAACAATAGAGAGAAACATCAATTGCAATTTGATGGTCTGATCTT 1032
Qy 1230 ----- 1229
Db 1033 CTGAGAGATGAACAGCCAGGACTGGTGATAGTCAACCGCAGACCACTTAAGTAGC 1092
Qy 1230 -----TGGATGTAGTGGGAAAGATGGGAATGCTGCTATGACACCGCCATC 1272
Db 1093 TTCTGCACTGGCAGTCTGATTTGATGGGAAAGATGGGAATGCTGCTATGACACCGCCATC 1152
Qy 1273 TTACTGACTGCTGATATATTTCTCGAAGATAGAGCTCTGACACTTTGGGATTT 1332
Db 1153 TTACTGACTGCTGATATATTTCTCGAAGATAGAGCTCTGACACTTTGGGATTT 1212
Qy 1333 GCACCCATTAAGTGAATGTGTAGTAAATATCGACCTGACGATTAATGAAGATACAGGT 1392
Db 1213 GCACCCATTAAGTGAATGTGTAGTAAATATCGACCTGACGATTAATGAAGATACAGGT 1272
Qy 1393 CTTGGAGCTGCTTACCACTGCTGAGTCTGAGACACAACTTTGGCATGATTCATGAT 1452
Db 1273 CTTGGAGCTGCTTACCACTGCTGAGTCTGAGACACAACTTTGGCATGATTCATGAT 1332
Qy 1453 GAGAGAGGAACTGTGTTAAAGTCCGAGGCAACATCATGTCCCTTACATTTGGCAGGA 1512
Db 1333 GAGAGAGGAACTGTGTTAAAGTCCGAGGCAACATCATGTCCCTTACATTTGGCAGGA 1392
Qy 1513 CGCAATGAGTCTTCTCTGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1572

Db 1393 CGCAATGAGTCTTCTCTCTGCTCACCCTGCAGCGCCAGTATCTACACAAATTTCTAAGC 1452
Qy 1573 ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAGGAATATCAAGATATCCT 1632
Db 1453 ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAGGAATATCAAGATATCCT 1512
Qy 1633 GAGAAATGCGCAGAGAGATTAATGATGCAAAACACACAGTGGAGTTCGAGAG 1692
Db 1513 GAGAAATGCGCAGAGAGATTAATGATGCAAAACACACAGTGGAGTTCGAGAG 1572
Qy 1693 AAAGCAAGCTCTGCTGCTGAGCTTTAAAGAGCAATCTGTAAAGCCTGTGGTGGCCAT 1752
Db 1573 AAAGCAAGCTCTGCTGCTGAGCTTTAAAGAGCAATCTGTAAAGCCTGTGGTGGCCAT 1632
Qy 1753 CGTATTTGAGAGAAATGTGAGACTTAAATTTATGCCAGCAGACAGGCAATTTGTGG 1812
Db 1633 CGTATTTGAGAGAAATGTGAGACTTAAATTTATGCCAGCAGACAGGCAATTTGTGG 1692
Qy 1813 CATGATGCTGCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1872
Db 1693 CATGATGCTGCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1752
Qy 1873 ACCCATGCTGCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1932
Db 1753 ACCCATGCTGCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1812
Qy 1933 GGAGTATCTCATAGAGTCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1992
Db 1813 GGAGTATCTCATAGAGTCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1872
Qy 1993 TGTGAGGCTCCACTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 2052
Db 1873 TGTGAGGCTCCACTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1932
Qy 2053 GTTGAATCTTCTGCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 2112
Db 1933 GTTGAATCTTCTGCTGCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 1992
Qy 2113 TACAGTGGAGCTTACACTCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 2172
Db 1993 TACAGTGGAGCTTACACTCTGCTGAGCTGTGTGAATATGCTGATGAAGCCCCCAAGCC 2052
Qy 2173 CGAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2232
Db 2053 CGAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2112
Qy 2233 TCGGAGGATAGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2292
Db 2113 TCGGAGGATAGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2172
Qy 2293 GTCTTTGATCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2352
Db 2173 GTCTTTGATCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2232
Qy 2353 TGCAAGATTCAGAGGCTCTCAACCAAGCACCACCAACCAACCAACCAACCAACCAACCAACCA 2412
Db 2233 TGCAAGATTCAGAGGCTCTCAACCAAGCACCACCAACCAACCAACCAACCAACCAACCAACCA 2292
Qy 2413 GTCAACATTTCTTCTGAGGCTGAGTATCGCATCTATGAAATGAACGCTCTTACCTCC 2472
Db 2293 GTCAACATTTCTTCTGAGGCTGAGTATCGCATCTATGAAATGAACGCTCTTACCTCC 2352
Qy 2473 TACATTTCTGTCGAATGCTTCTGAGGCTGAGTATCGCATCTATGAAATGAACGCTCTTACCTCC 2532
Db 2353 TACATTTCTGTCGAATGCTTCTGAGGCTGAGTATCGCATCTATGAAATGAACGCTCTTACCTCC 2412
Qy 2533 TGGCCCGCCCGGTACAAATTTTGGGCACTACTTTCGACTACAGAGCTCTTATATATGAG 2592
Db 2413 TGGCCCGCCCGGTACAAATTTTGGGCACTACTTTCGACTACAGAGCTCTTATATATGAG 2472
Qy 2593 CCGGAGAACTTAATTCGCTGCTGAGCACAACCAACGAGACACTGATTTGTGAGCTCTTGTGTTT 2652
Db 2473 CCGGAGAACTTAATTCGCTGCTGAGCACAACCAACGAGACACTGATTTGTGAGCTCTTGTGTTT 2532


```

QY 860 ACTTATGCAAACTCTACTGTATCGCAGAAAGGATTGATTTCTTCTTTCTTTGTCAAATA 919
Db 625 ACAGAGCCCATCTCTCTGGGCGAGTGAGTCTCTGGTCACTC-----AAGG 672
QY 920 AAGTCAAGATGGGACTCATCTGTCGGAGGATAGCGTAATGTTTGTATAGATGGGATAT 979
Db 673 ACATGGGAGTGGCACAATCAACCCCTGCAACAGCAGCGACTTCGCTGGGACTGCGCAA 732
QY 980 GTGAGCTCAGTGTGGTGTCAATCTGCGC-ACATGCCCCAGCCTCCCAAGAGAACCTC 1038
Db 733 AACGACGATTTCTGTGGAAGAGCGCAAGAAATACATGCCCCAGCCTCCCAAGAGAACCTC 792
QY 1039 TTCACTTGTCCAGATGAGTATAGTCTTGTCTTACGGCATAAAGCGTCTCTTCTGAGTCC 1098
Db 793 TTCACTTGTCCAGATGAGTATAGTCTTGTCTTACGGCATAAAGCGTCTCTTCTGAGTCC 852
QY 1099 CATAGAAATGAAGAACTGACGTGGAGACCTTGGTGGTGTGCGACAAAAGATGACAA 1158
Db 853 CATAGAAATGAAGAACTGACGTGGAGACCTTGGTGGTGTGCGACAAAAGATGACAA 912
QY 1159 AACCATGGCCATGAAATATCACCACTACGTCTCAAGATCTCAACATGATGTATCTCT 1218
Db 913 AACCATGGCCATGAAATATCACCACTACGTCTCAAGATCTCAACATGATGTATCTCT 972
QY 1219 TTATTCAGAGA----- 1229
Db 973 TTATTCAGAGA----- 1229
QY 1230 ----- 1239
Db 1033 CTAGAGATGAACAGCAGGACTGTGTGATAAGTACCCAGCAGACCAACCTTAAGTAGC 1092
QY 1230 -----TGGATTGATGGGAAGATGGGACTCGTCATGACCAACCCATC 1272
Db 1093 TTCTGCCAGTGGCAGTCTGGATTGATGGGAAGATGGGACTCGTCATGACCAACCCATC 1152
QY 1273 TTAAGTACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCTGTGACACTTTGGGATTT 1332
Db 1153 TTAAGTACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCTGTGACACTTTGGGATTT 1212
QY 1333 GCACCCATAGTGGATGTAGTAAATATTCGGAGCTGCAAGATTAAGTAAAGATACAGGT 1392
Db 1213 GCACCCATAGTGGATGTAGTAAATATTCGGAGCTGCAAGATTAAGTAAAGATACAGGT 1272
QY 1393 CTGGAGCTGGCTTACCACTTCCCATGAGTCTGGACACAACCTTTGGCAGATGATCATGAT 1452
Db 1273 CTGGAGCTGGCTTACCACTTCCCATGAGTCTGGACACAACCTTTGGCAGATGATCATGAT 1332
QY 1453 GGAAGAGGAACATGTGTAAAGTCTCGAGGGGCAACATCATGTCCCTACATGTCGAGGA 1512
Db 1333 GGAAGAGGAACATGTGTAAAGTCTCGAGGGGCAACATCATGTCCCTACATGTCGAGGA 1392
QY 1513 CGCAATGGAGTCTTCTCTGTCACCTGCGAGCGGCGAGTATCTACACAAATTTCTAAGC 1572
Db 1393 CGCAATGGAGTCTTCTCTGTCACCTGCGAGCGGCGAGTATCTACACAAATTTCTAAGC 1452
QY 1573 ACCGCTCAA 1581
Db 1453 TCAGTGAAG 1461

```

RESULT 4

```

US-09-392-184-5/c
; Sequence 5, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PROTEASE HOMOLOGS
; FILE REFERENCE: 5800-55
; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09

```

```

; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 5357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(5357)
; OTHER INFORMATION: reprotolysin (ADAM family of metalloprotease)
; NAME/KEY: misc_feature
; LOCATION: (1)...(5357)
; OTHER INFORMATION: n = A,T,C or G
; US-09-392-184-5

Query Match 7.1%; Score 204.6; DB 4; Length 5357;
Best Local Similarity 50.9%; Pred. No. 2.4e-48;
Matches 787; Conservative 0; Mismatches 714; Indels 44; Gaps 11;

QY 1258 CATGACCAAGCCATCTACTGCTGTGGATATATGTTCTGGAGAAATAGGCTGT 1317
Db 4316 CATGACACTGCCATCTGCTCACAGAAAGGACCTGTGTGACGACCATGAACCGGCCCTGT 4257
QY 1318 GACACTTTGGGATTTGCACCCATAAGTGAATGTGTAGTAAATATCGCAGCTGCCAGATT 1377
Db 4256 GAGACCTTGGGATTTGCACCCATAAGTGAATGTGTAGTAAATATCGCAGCTGCCAGATT 4197
QY 1378 AATGAAGATACAGCTTTGACATGCGCTTTCACCATTTGCCATGCGAGTCTGGACACACTTT 1437
Db 4196 AAGCGAGACACGGGCTTGGCTGGCTTTCACTGTAGCCCAACGAGCTCGGGCACAGTTT 4137
QY 1438 GGATGATTCATGATGGAGAGGGAACATGTGTAAAGATGCCAG-----GGCAACATC 1491
Db 4136 GGATGATTCATGATGGAGAGGGAACATGTGTAAAGATGCCAG-----GGCAACATC 4077
QY 1492 ATGTCCCTTACATTTGGCAGACGCAATGGAGTCTTCTCTGTGTACCTCTGTCGAGCCGAG 1551
Db 4076 ATGTCTCCACAGCTCTGTACGACGCGCTCCCTCACCTGGTCTCCGCTCGAGCCGCCAG 4017
QY 1552 TATCTACACAAATTTTAAAGCAGCCGCTCAAGCTATCTGCTTGTGCTGTGATCAGCAAGCTT 1611
Db 4016 TATATCACCAAGTTTCTTGAACCTTGGTGGGCGCTGTGCTGTGACGACCTCC---TGCC 3960
QY 1612 GTGAAGGAATPACAAATGATCTGAGAAATTCGCCAGGAGAAATATATGATCAACACACAG 1671
Db 3959 AAGGACATTTATCGACTTCCCTCGGTGCCACCTGTCTCTATGATGATGAGCCACCCAG 3900
QY 1672 TGCAGTGGCAGTTGGAGAGAAAGCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1731
Db 3899 TGCAGCTTCAGTACCGGGGCTTACTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3846
QY 1732 TGTAAAGCCCTGTGGTGGCCATCGTATTGGAAGGAAATGTGAGACTAAATTTATGCGCAGA 1791
Db 3845 TGCACACACTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3788
QY 1792 GCAAGAGGCAATTTGTGGGAGTATGATGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1851
Db 3787 GTGACGGCACCCGCTGTGGGAGAAATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3731
QY 1852 GGTGATGAAGGCCCCCAAGCCCTTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1911
Db 3730 GTGGCTTCCGGCCCGAGGCGGTGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3671
QY 1912 TGCTTCAAGGACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1971
Db 3670 TGCTTCAAGGACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3611
QY 1972 CCATCGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2031
Db 3610 CCCAAATACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3551
QY 2032 CAGAAATGTCCCCGGGACAGTGTGTGACTTTCCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2091

```


Db 3550 CAGGCTGCTGCTGGCGCCCTCTTCCGCCACGCTCCAGTGCAGCCACTTTGACGCT 3491
 Qy 2092 AGACGATTCAGAGGGCGGCACCTACAGTGGAGCCTTACACTCAAGTAGAGATCAGGAC 2151
 Db 3490 ATGCTCTACAGGGCGAGCTGCACACATGGGTGC-----CGTGGTCAATGAGTGAAC 3437
 Qy 2152 TTATGCAAACTCTACTGTATGCGAAGAGATTGATTCTTTCTTTCTTGTCAATAAA 2211
 Db 3436 CCGTGGAGCTGCTGCGCGCGCGCGAATGATGATCTTTGCGGAGAGCTCGGAGCGCC 3377
 Qy 2212 GTCAGAGATGGAGCTCCATGCTCGAGGAT-----AGCGTAAATGTTGTATAGATGG 2265
 Db 3376 GTGGTCAATGGACACCCCTCTCTACAGTCCAGGTCAGCGGACCTCTGATCAACGGC 3317
 Qy 2266 ATATGTGAGAGATTGGATGTGACAATGTCCTTGGATCTGATGCTGTGTTGAAGAGCTGT 2325
 Db 3316 ATCTGTAAAGAACGTGGGCTGTGACTTGGAGATTGACTCGGTGCTATGGAGGACCGTGT 3257
 Qy 2326 GGGTCTGTAAACGGGAATAACTCAGCTGSCAGATTACAGGGGTCTCTACACCAAGCAC 2385
 Db 3256 GGTGTGTGCCACGGCAACCGCTCCACCTGCCACACGCTGAGCGGACCTTCG---AGAG 3200
 Qy 2386 CACCACCAACACAGPATTATACATGTCACATTCCTTCTGAGCCCGGAGTACCGC 2445
 Db 3199 GCGAGGGCTGGGGTATGTGGATGTGGGCTGATCCACGCGGCGCACGCGAGATCCG 3140
 Qy 2446 ATCTATGAATGAAGTCTCTACCTCCTACATTTCTGTGGCA---ATGCCCTCAGAGG 2502
 Db 3139 ATCCAGAGGTTGCCGAGGCTGCCAACTTCTGGCACTGCGAGGTGAGGACCCGGAGAG 3080
 Qy 2503 TACTACTGAATGGGCACTGGACCGCTGGACTGGCGCGCGCGGTACAAAATTTTCGGGCACT 2562
 Db 3079 TACTTCTCAATGGTGGCTGGACCATCCAGTGAACGGGACTACAGGTGGCAGGACC 3020
 Qy 2563 ACTTCGACTACAGAGCTCTATATGAGCCCGGAGAGACTTAATGCTACTTGGACCAACC 2622
 Db 3019 ACCTTCACATACGACGACGGGGCAACTGG---GAGAACCTCAGCTCCCGGGTCCCAACC 2963
 Qy 2623 AACGAGACTGATTTGGAGCTGCTGTTTTCAGGGAAGAACCCGGGTGTTGCTGGGAA 2682
 Db 2962 AAGAGCTGTCTGGATCAGCTGCTGTTTTCAGGAGAGACMACCTTGGGTGCACTACGAG 2903
 Qy 2683 TACTCAGCTGCTGGTGGGACCGGAGAGAGAGCCCTGCGCCAGCCAGCTACACTTGG 2742
 Db 2902 TACACCATCCAGAGGAGGAGGTGGCCACGACGAGGTCCCGCGCGCGTGTCTCTCTGG 2843
 Qy 2743 GCCATCGTGGCT---CTGAGTCTCCGTGCTGCGGAGGGGT 2784
 Db 2842 CATTATGGGCGCTGGACCAAGTGCACAGTCACTTGGCGCAGAGGT 2798

RESULT 5

US-09-369-364A-6
 ; Sequence 6, Application US/09369364A
 ; Patent No. 6391610
 ; GENERAL INFORMATION:
 ; APPLICANT: Apce, Suneel
 ; APPLICANT: Hurskainen, Tiina L.
 ; APPLICANT: Hirohata, Satoshi
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
 ; FILE REFERENCE: 26473/4007/10-30-00
 ; CURRENT APPLICATION NUMBER: US/09/369,364A
 ; CURRENT FILING DATE: 1999-08-06
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 3218
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens ADAMTS-7
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (13)..(3003)
 US-09-369-364A-6

Query Match 6.3%; Score 182.6; DB 4; Length 3218;
 Best Local Similarity 50.8%; Pred. No. 3.8e-42;
 Matches 714; Conservative 0; Mismatches 649; Indels 42; Gaps 10;
 Qy 1258 CATGACCAGCCCATCTTACTGACTGCTGTGATATATGTTCTTGGAGAAATGAGCCCTGT 1317
 Db 1015 CATGACACTGCCATCTCTCACCAGAAAGGACCTGTGTGAGCCATGAACCGGCCCTGT 1074
 Qy 1318 GACACTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATT 1377
 Db 1075 GAGACCTTGGGACTGTCCCATGTGCGGCGCATGTGCCAGCCGACCGCAGCTGCAGCATC 1134
 Qy 1378 AATGAAGATACAGGTCTTGGACTGCGCTTACCATTTGCCATGTGCCATGTGTCGACACACTTT 1437
 Db 1135 AACGAGGACACGGGCTGCGCGTGGCTTCTACTGTAGCCACGAGCTCGGCGCAGACTTTT 1194
 Qy 1438 GGCATGATTTCATGATGGAGAGGGAACATGTGTAAAGTCCGAG-----GGCAACATC 1491
 Db 1195 GGCATTCACANTGAACGGAAGCGGCAATGACTGTGAGCCGTTGGAAACGACCTTTTCATC 1254
 Qy 1492 ATGTCCCTTACATTTGGCAGGACGCAATGGAGTCTTTCTCTGTGTACCTGTCAGCGCCAG 1551
 Db 1255 ATGTCTCCACAGTCTCTGTACGACCGCTCCCTCTACCTGTGTCGCTGCGCGCAGC 1314
 Qy 1552 TATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGATGATGACCAAGCCT 1611
 Db 1315 TATATCACCAGGTTCTTGACCGTGGTGGGCTGTGCTGGAGCAGCCCTCT---GCC 1371
 Qy 1612 GTGAAGGAATACAAATATCTCTGAGAAATTTGCCAGAGAAATATATATGATGCAAAACACACG 1671
 Db 1372 AAGGACATTTATGACTTCCCTCTCGTCCACCTCGGCTCTCTATGATGTAAAGCACACG 1431
 Qy 1672 TGCAGTGGCAGTTGGAGAGAAACCAAGCTCTGCTGCTGCTGAGACTTTAAAGAGACATC 1731
 Db 1432 TGCCTGCTCCAGTACGGGCTTACTCTGCTTTCTGC-----GAGGACATGGATAATGTC 1485
 Qy 1732 TGTAAAGCCCTGTGGTGCCATCTGTATTTGAAGAAATGTGAGACTAAATTTATGCGACGA 1791
 Db 1486 TGCCACACACTCTGTGCTCTG---TGGGACCACTGTCTACTCCAGCTGGATGACGCT 1542
 Qy 1792 GCAAGAGGACAAATTTGTGGGATGACATGTGCTGCGGGAGGAGACATGTGTGAATAT 1851
 Db 1543 GTGAGCGGACCCGCTGTGGGAGAAATAGTGTCTCTAGTGGGAGTGGCT---ACCC 1599
 Qy 1852 GGTGATGAAGGCCCAAGCCACCATGSCCACTGGTGGAGTGTCTTCTGTGTCGCCA 1911
 Db 1600 GTGGGCTTCCGGCCCGAGGCCGTGGATGGTGGCTGTGCTGGAGCGCTGGTCCATC 1659
 Qy 1912 TGCTCCAGGACCTGCGAGGCGGAGTATCTCATAGAGTGGCTCTGACCAACACCCCAAG 1971
 Db 1660 TGCTCAGGAGCTGTGGCATGGGCGGTACAGAGCGCGGAGCGGAGTGCACGACGCTACG 1719
 Qy 1972 CCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTCTGCAACAGT 2031
 Db 1720 CCCAATACAAAGGCGAGTACTGTGTGGTGGGCGCAAGCGCTTCCGCTCTGCAACCTG 1779
 Qy 2032 CAGAAATGTCCCGGAGCAGTGTGTGACTTCCGTGCTGTCTGATGTGCGGAGCAACAGC 2091
 Db 1780 CAGGCTTGGCTCTGCTGGCGGCCCTCTTCCGCCACGCTCCAGTGCAGGCACTTTGACGCT 1839
 Qy 2092 AGACGATTACAGAGGGCGGCACTACAGTGGAGCCCTTACACTCAAGTAGAGATCAGGAC 2151
 Db 1840 ATGCTCTACAGGGCCAGCTGCACACATGGGTGC-----CCGTGGTCAATGAGTGAAC 1893
 Qy 2152 TTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTTGTCAATAAA 2211
 Db 1894 CCGTGGAGCTGCACTGCGCGGCCCGGAATGAGTACTTTGCCAAGAGCTGCGGAGCGCC 1953
 Qy 2212 GTCAGAGATGGAGCTCCATGCTCGAG-----GATAGCCGTAATGTTGTATAGATGG 2265
 Db 1954 TGTGTCGATGGCACCCCTCTCTACCGAGGTCGAGCGGAGCCTCTGATCAACAGGC 2013

QY 2482 GTGGCAATGCTCAGAGGTAAGTACTGATGGGCACTGGACGGTGGAGTGGCGCCG 2541
Db 2173 CTGAAGGGAGACGAGAGTCCCTGCTGGAGGGGCTGCTGGAGCCCGCCAGCCAC 2232
QY 2542 CGGTACAAATTTTCGGGCACTACTTTTCGACTACAGACGGTCTCTATATAGCCCGAGAAC 2601
Db 2233 CGTCTGCTCTAGCTGGGACCACTTCACTGCGACAGGGGCGCCAGAGCTCCAGAGC 2292
QY 2602 TTAATCGCTACTGACCAACCAACGACACACTGATTTGGAGCTGCTG 2649
Db 2293 CTGGAAGCCCTGGGACCGGATTAATGATCTCTCATGCTGATGGTGGCTG 2340

RESULT 7

US-09-369-364A-4
; Sequence 4, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22)..(2601)
; NAME/KEY: misc feature
; LOCATION: (1365)
; OTHER INFORMATION: n = C
; NAME/KEY: misc feature
; LOCATION: (1620)
; OTHER INFORMATION: n = C
US-09-369-364A-4

Query Match 5.0%; Score 144; DB 4; Length 2848;
Best Local Similarity 48.5%; Pred. No. 4.5e-31;
Matches 464; Conservative 0; Mismatches 481; Indels 12; Gaps 2;

QY 1726 GACATCTGTAAGCCCTGGTGCCATCTGATGGAGGAATGTGAGACTAAATTTATG 1785
Db 1390 GAAGTGTGTAGAGCTCTGGTGTCTCAGCAAAAGCAACCGTGTGTCCACCAAGATTT 1449
QY 1786 CCAGCAGCAGAGGACAAATTTGTGGGCATGACATGTGTGCGGGGAGGACAGTGTGTG 1845
Db 1450 CCAGCAGCTGAGGGGACACTGTGTCAAACTGGGAATATTGAAAAGGGTGTGTATCAG 1509
QY 1846 AATATGTGTGAT-----GAGGCCCAACCCACCCATGGCCACTGTCGACATGG 1896
Db 1510 GAGATGTGTGTTCCCTTTTGGCACTTGGCCCGAGCATGATGATGGGGTGGGGTCCCTGG 1569
QY 1897 TCTTTCTGTTCCCATCTCCAGGACCTGCGAGGGGGAGTATCTCATAGGAGTGGCCCTC 1956
Db 1570 TCACTATGGGAGAGTGCAGCAGGACCTTGGGGGGAGGGTGTCTCATCCCTCAAGACAC 1629
QY 1957 TGACACACCCCAAGCATGCTATGAGGAGATCTCTGTGAGGGTCCCTGCGACTCTG 2016
Db 1630 TGTGACAGTCCAGCACCTTCCGAGGTGGAAAATATTGCTTTGGGAAAGGAAACCGTAT 1689
QY 2017 AAGCTCTGCAACAGTCAAGAAATGTCGCCGAGCAGTGTGACTTCCGTGCTCTCAGTGT 2076
Db 1690 CGCTCCGTGTAACAGATCATCCCTTTGGGTTCGAGATTTTCGAGAGAAACAGTGT 1749
QY 2077 GCGGACACACAGCAGACAGTTCAGAGGGGCGGCACTACAAGTGGAGCTTTACACTCAA 2136
Db 1750 GCAGACTTTGACAATATGCGCTTTCCGAGGAAAGATATTATACTGGAACCCCTACTACT 1806

QY 2137 GTAGAGATCAGGACTTATGCAAACTCTACTGCTATCGCAGAGGATTTGATTTCTCTTT 2196
Db 1807 GAGGTGGGGTAAACCTTTGTCATTTAACTGCTTGGCTGAGGTATTAATTTCTACTACT 1866
QY 2197 TCTTTGTCAAATAAAGTCAAGATGGGACTTCATCTCGGAGGATPAGCCGTAATGTTGT 2256
Db 1867 GAACGTGCTCCTCGGTGATCGATGGACCCAGTSCAATGCGGATTCACATCGATATCTGC 1926
QY 2257 ATAGATGGATATGTGAGAGAGTTGGATGTGACATGTCCTTGGATCTGATGCTGCTGAA 2316
Db 1927 ATCAATGGAGATGCAAGCACGTGCTGTGATAATATTTTGGATCTGATGCTAGGAA 1986
QY 2317 GACGTCTGTGGGTGTGTAAACCGGAAATAACTCAGCTGCGACGATTCACAGGGGTCTCTAC 2376
Db 1987 GATAGATGTCGAGTCTGTGGAGGGCGGAGCACATGTGATGCCATTTGAAGGGTCTTTC 2046
QY 2377 ACCAGCACACACACACACAGTATTATACATGCTCACCATTCTCTTGGAGCCCG 2436
Db 2047 AATGATTCATCTCCAGGGAGGCTACATGGAAGTGGTGCAGATACCAAGAGGCTCTGTT 2106
QY 2437 AGTATCCGCATCTATGAAATGAACGTCTCTACTCTTACTCTTCTGTGCGCAATGCCCTC 2496
Db 2107 CACATTCAGTGTAGAGAAGTTCCTCATGTCAGAGAACTATATTGCTTTAAATCTGAAG 2166
QY 2497 AGAAGTACTACTGATGGGCACTGGACCTGGAGCTGCGCCGCGGCTGACAAATTTTCG 2556
Db 2167 GATGATTACTATATTAATGCTGCTGAGCTATTTGACTGCGCTAGGAAATTTGATGTTGCT 2226
QY 2557 GGCACCTACTTTCGACTACAGACGGTCTTATATATGAGCCCGAGAACTTATCGCTACTGGA 2616
Db 2227 GGGACAGCTTTTCATTTACAGAGACCACTGATGAACAGAACTCTTGAAGCTCTAGT 2286
QY 2617 CCAACCAAGCAGACACTGATGTGGAGCTGTGCTGTTTCAGGAGGAAACCCCGGTGTT 2673
Db 2287 CCTACCTCGAATAATCTCATGCTGATGCTGCTCTGCTTCAAGAACAGAAATTTGGGAAT 2343

RESULT 8

US-09-491-522-2
; Sequence 2, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Lapiere, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-493-4935
 TELEFAX: 650-493-5556
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2450 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-491-522-2

Query Match 4.1%; Score 118.4; DB 4; Length 2450;
 Best Local Similarity 48.0%; Pred. No. 9.6e-24;
 Matches 573; Conservative 0; Mismatches 591; Indels 30; Gaps 7;

QY 1308 TGAGCCCTGTGACACTTTGGGATTTGACACCATAGTGGAAATGTAGTAAATATGCGAG 1367
 DB 1194 TGGGCTTCCGGCATGCAAGCTATGCTCTGTGACCGCATGTGCCATCCGGTCCGAG 1253
 QY 1368 CTGCACGATTAATGAAATACAGGTCTTGGACTGGCTTCCACATTCGCCATGAGTCTGG 1427
 DB 1254 CTGCACCTGTAACCATGAGGAGGCTTCTCTCAGCGTCTTGTGGTGGCCCATGAGACTGG 1313
 QY 1428 ACACAACTTTGGCATGATTCATGATGGAGAGGAAACATGTGTAAAGTCCGA----- 1481
 DB 1314 CCAGCTGTGGCATGGAGCAGCAGCGGCGGCAACCGCTGTGGCAGAGGTGGCGCT 1373
 QY 1482 GGGCAACATCATGCTCCCTACATTTGGCAGGACCAATGGAGTCTTCTCTGGTCACTGG 1541
 DB 1374 GGGCAGCATATGCGGCCCTGGTGGAGGCGCTTCCACCGTGTCCACTGGTCCGCGCTG 1433
 QY 1542 CAGCGCCAGTATCTACACAAATTTCTAAGCACCCTCAAGCTATCTGCGCTTGTGATCA 1601
 DB 1434 CAGCAGCAGAGCTGAGCGCTACCTGCACTCTCTA-----CTGCTGTGGATGA 1487
 QY 1602 GCCAAGCCTGTGAAGGATACAGTATCTCTGAGAAATTTGCGAGGAAATATATGATGC 1661
 DB 1488 CCGCTTCCGCC---ACGACTGGCGGCGCTCCCGAGCTCCCGGACTGCACTACTCCAT 1544
 QY 1662 AAACACACAGTGCAGTGGCAGTTCGGAGAGAAAGCAAGCTCTGCACTGTGCACTTAA 1721
 DB 1545 GACAGAGATGCGCCTTGACTTGGCTGGCTGGCTACATGATGTCAGCGGTTCGGAC 1604
 QY 1722 AAGGACATCTGTAAGCCCTGTGGTGCATCTGATTTGGAAG-----AAATGTGAGAC 1775
 DB 1605 CTTTGACCCCTGCAAGAGCTGTGGTGCAGCATCTCTGACACCCCTACTTTTCAAGAC 1664
 QY 1776 TAAATTTATGCCAGCAGCAGAGGACAAATTTGGGATGACATGTTGGTGGCGGAGG 1835
 DB 1665 CAGAGAGGGGCCCCCTTGGAGCGACATATGTGTGACCTGGCAAGCATGTTTAAAG 1724
 QY 1836 ACAGTGTGAAATATGATGAGGCCCCCAAGCCCATGCGCATGCGCACTGTGCGACTG 1895
 DB 1725 ACATGCTATCTGGCTGACACCTGACATCTCTCAACAGGGAC---GGCAGCTGGGGCGCTG 1781
 QY 1896 GTCTTCTTGGTCCCATGCTCCAGGACTCGCGAGGGGAGTATCTCATAGGAGTGGCCT 1955
 DB 1782 GAGTCCGTTGGTCTCTGCTACGTACTGTGTCAGCGCGTGAAGTTCAGGACCCCA 1841
 QY 1956 CTGCACCAACCCCAAGCATCGCATGAGGAGGATTTGTGAGGGCTCCACTGCGACTCT 2015
 DB 1842 GTGTGACAAACCCACACCGGCGGCAAGCGGGCGGCACTGTCTGGGCTTGCCTACGACTT 1901
 QY 2016 GAAGCTCTGACAGTCAAGATGTCCTGGGAGAGTGTGACCTTCCGTCTGCTCAGTG 2075
 DB 1902 CCAGCTCTGACGCGCCAGGACTGCCCGACTCTCTGCTGACTTCCGAGGAGAGTG 1961
 QY 2076 TGCCGA---GCACAAACAGCAGCAGATTCAGAGGGCGGCACTACAGTGGAGGCTTACAC 2132
 DB 1962 CCGCCAGTGGGACCTGTATTCAGACACGGCGAGCGCCAGCACCCTGCTGGTCCCGCAGA 2021
 QY 2133 TCAAGTAGAGATCAGGACTATGCAACTCTACTGATTCGACAGAGATTTGATTTCTT 2192

DB 2022 GCACCGGATGCCAAGGAGAGATGCCACCTGTACTGCGAGTCCAGGAGACCGGGAGGT 2081
 QY 2193 CTTTCTTTTGCATAAATCAAGATGAGGACTCCATGCTC---GGAGATAGCGTAA 2249
 DB 2082 GGTGTCATGAGCGCATGTTGATGATGAGCGGCTGCTCTACAGGACCGCTTCA 2141
 QY 2250 TGTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAATGTCCTTGGATCTGATGC 2309
 DB 2142 CTTCTGTGTGCGGGGACTGACAGAGAGTGGGCTGTGACGGTGTGATCGGCTCCAGCAA 2201
 QY 2310 TGTGAGAGCTGTGCGGGTGTGTAACGGGAAATCACTCAGCTGCACGATTCACAGGG 2369
 DB 2202 GCAGAGAGACAGTGTGGCGTGTGCGAGGGGACACAGCCCTGCAAAAGTGGTCHAGG 2261
 QY 2370 TCTCTACACCAAGCACCACACACCAACCAAGTATTATCATAGTGCATCCATTCCTTCTGG 2429
 DB 2262 CACGTTACACAGGTCACCCAAAGCATGTTTATCATCAAGATGTTTGTAGATCCCTGAGG 2321
 QY 2430 AGCCCGAGTATCCGATCTATGAAATGAAGTCTTACCTCTACATTTCTGT 2483
 DB 2322 ACCCAGACACTGCTCTATTGAGGAGGTAGACGCCACACCATCTGGCGCT 2375

RESULT 9

US-09-491-522-1
 ; Sequence 1, Application US/09491522
 ; Patent No. 6428998
 ; GENERAL INFORMATION:
 ; APPLICANT: Collige, Alain
 ; APPLICANT: Lapiere, Charles M.
 ; APPLICANT: Prockop, Darwin J.
 ; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE.
 ; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds, LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2811
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FastSeq for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/491,522
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/886,333
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Abrams, Samuel B.
 ; REGISTRATION NUMBER: 30,605
 ; REFERENCE/SOCKET NUMBER: 8389-0060-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-493-4935
 ; TELEFAX: 650-493-5556
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6592 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-491-522-1

Query Match 4.1%; Score 118.4; DB 4; Length 6592;
 Best Local Similarity 48.0%; Pred. No. 1.7e-23;
 Matches 573; Conservative 0; Mismatches 591; Indels 30; Gaps 7;

QY	1543	AGCGCCAGTATCTACACAAATTTCTAAGACCGGOTCAAAGTATCTGCGCTTGCATGACG 1602
DB	1473	AGCCAGCAGGAGTGAGCGCTACCTTGCACTCTCTATGA-----CTGCGCTGGGGATGAC 1526
QY	1603	CCAAAGCCCTGTGAGGAATACAAATATCTGAGAAATTCGCAAGGAAATTTATATGATGCA 1662
DB	1527	CCCTTCAC--CCACCACTGGCGCGCTGCCCGAGCTCCCGGGCTGCACACTACTCCATG 1583
QY	1663	AACACACAGTGCAGGTGGCAGTTTCGAGAGAAAGCCAAAGCTCTGCACTGTGCACTTTAAA 1722
DB	1584	AACGACAGTGCCTTTTCGACTTTTGCCTTGGTTACATGATGTGATCCCGTTCGGACC 1643
QY	1723	AAGGACATCTGTAAGCCCTCTGGTGGCATCTGATTGGAA-----GGAAATGTGAGACT 1776
DB	1644	TTCCGACCGGTGCAAAAGAGCTGTGGTGAGCCACCCCTGACAAACCCCTACTTTTTCGAAGACA 1703
QY	1777	AAATTTATGCCACAGCAGAAAGGACAAATTTGTGGCATGCAATGTGTGGTGGCGGGAGGA 1836
DB	1704	AAGAAGGGGCCACCTTGGATGGACCAATGTGTGGCTTGGCAAGCACTGCTTTAAAGGA 1763
QY	1837	CAGTGTGTGAATATATGGTGAATGAAGGCCCCCAAGCCCAACCATGCGCACTGTGTCGACTGG 1896
DB	1764	CACTGCATCTGGCTGACACCTGACATCTCAAAAC---GAGATGGCAACTGGGGTGCTCG 1820
QY	1897	TCTTCTTGTGTCCTCATGCTCCAGAGCTCGGAGGGGGAGTATCTCATAGGAGTCCGCTC 1956
DB	1821	AGTCCCTTGGCTCTCTGCTCGGTACTCTCGGCACAGGTGTGAAGTTCAAGACCCGTGAC 1880
QY	1957	TGCACAAACCCCAAGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCACTGCGCACTCTG 2016
DB	1881	TGGCAACCCCAACCCAGCAATGGGGCCGCACATGCTCGGGCCCTCGCCTACGATTTTC 1940
QY	2017	AAGCTCTGCACAGTACAGAAATGTCCCGGGACAGTTTGACACTTCGGTCTGCTCAGTGT 2076
DB	1941	CAGCTCTGCAACTTCGACAGACTCCCTGTAGCGCTGGCCGACTTCGCGAGGAGCACTGC 2000
QY	2077	GCCGA---GCACAACAGCAGACGATTACAGAGGGCGCACTACAAGTGGGAAGCCTTACACT 2133
DB	2001	CGGCAGTGGACCTGTACTTTCGAGCATGTGACGCTCAACACCCCTTGGCTGCCCCACGAG 2060
QY	2134	CAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAAGGATTTGATTTCTTC 2193
DB	2061	CACCGGAGCCCAAGGAGCGGTGTCTCTCTACTGTAGTCCAAAGGAGACCGGGAGGTG 2120
QY	2194	TTTTCTTTGTCAATAAAGTCAAAGATGGGACTCCATG---CTCGGAGGATAGCCGTAAT 2250
DB	2121	GTGTCAATGAACGTTATGGTGCATGACGGACACGCTGTCTTCTACAAAGACGCTTCAGC 2180
QY	2251	GTTTGTATAGATGGGATATGTGAGAGGTTGGATGTGACAAATCTCTTGATCTGATGCT 2310
DB	2181	CTCTGCGTGGCGACTCAGAAAGTGGGCTGTGACGGGTGATCGGCTCCAGCAAG 2240
QY	2311	GTTTGAAGCTCTGTGGGTGTCTAAGGAAATTAAGCTCAGCCTGACAGATTCAACGGGGT 2370
DB	2241	CAGGAGACCAAGTGTGTGTGTGGGAGGGGCAACTCCCACTGCAAGGTGTGCAAGGGC 2300
QY	2371	CTCTACCAACGACCAACCAACCAAGTATTATCAATGTGTCACTGTCCACTTCCTTCTGGA 2430
DB	2301	ACGTTCTCGCGCTCGCCCAAGAGCTTGGTTATCAAGATGTTTGNATCCCGGACGGA 2360
QY	2431	GCCCGGAGTATCCGCACTATGAAATGAACGTTCTTACCTTCCTACATTTCTGT 2483
DB	2361	GCCAGACACCTGTAAATCCAGGAAGACAGACACACCAAGCCATCACTTGGCCGT 2413

RESULT 11

RESOL 11
US-09-491-522-8

03-03-491-322-8
; Sequence 8, Application US/09491522

Sequence 3, Applicant
: Patent No. 6428998

FILE NO. 6428350
: GENERAL INFORMATION:

APPLICANT: Colique, Alain

APPLICANT: Lapierre, Charles M.

Db 1644 TTGACCCGTCGAAACAGCTGTGGTGGAGCCACCTGACACCCCTACTTTTGCAGACA 1703
QY 1777 AAATTTATCCAGCAGCAGAGACAAATTTTGGGCATGACATGTTGGCGGAGGA 1836
Db 1704 AAGAAGGGCCACCCCTGGATGGACCATGTGTGGCTGGCAAGCACTGCTTTAAAGGA 1763
QY 1837 CAGTGTGTGAATATGGTGTAGAGGCGCCCAAGCCACCCATGCGCTGCTCGACTGG 1896
Db 1764 CACTGCATCGGCTGACACCTTGACATCTCAAAAC---GAGATGGCAACTGGGTGGCTGG 1820
QY 1897 TCTTCTGTGTCCTCCATGCTCCAGGACCTCGGAGGGGGAGTATCTCATAGGAGTCGCCCTC 1956
Db 1821 AGTCCCTTCGGCTCTGCTGCGTACCTTGGGACAGGTGTGAAGTTCAAGACCCCTCAG 1880
QY 1957 TGACACACCCCAAGCATGCGATGAGAGGAATTTCTGTGAGGCTCCACTGCGACTCTG 2016
Db 1881 TGGGACAACCCACACCCAGCAATGGGGCGGCACATGCTCGGGCCCTCGCTACGATTTTC 1940
QY 2017 AGCTCTGCAACAGTCAGAAATGTCCCGGGACAGTGTGACTTTCGCTGCTCAGTGT 2076
Db 1941 CAGCTCTGCACTCGCAGGACTGCTGCTGACGCGCTGCGGCTGCGGAGGAGTGC 2000
QY 2077 GCGCA---GCACACAGCAGACGATTCAGAGGCGGCACCTACAAAGTGGAAAGCTTACACT 2133
Db 2001 CGGCAGTGGGACCTGTACTTCGAGCATGGTACGCTCAACACCACTGGCTGCCCCACGAG 2060
QY 2134 CAGTAGAGATCAGGACTTATGCAACTCTACTCTATGTCGAGAGGATTTGATTTCTTC 2193
Db 2061 CACCGGAGCCCAAGAGCGGTGTCTCTCTCTGTGAGTCCAAGGAGACCGGGAGGTG 2120
QY 2194 TTTTCTTTTTCAAATAAAGTCAAAGATGGGACTCCCATG---CTCGGAGGATAGCCGTAAT 2250
Db 2121 GTCTCCATGAAGGTATGTGTCATGCGGACAGCGCTGTTCTTACAGGAGCGCTTCAGC 2180
QY 2251 GTTTGTATAGATGGATATGAGAGAGTGGATGTGACATGTCTCTGATCTGATGCT 2310
Db 2181 CTCTGCTGCTGGGACTGACGAGAAAGTGGGCTGTGACGGGTGTATCGGCTCCAGCAAG 2240
QY 2311 GTTGAAGACGTGTGGGTGTGTAACGGGAATACTCAGCCGACGATTCACAGGGGT 2370
Db 2241 CAGAGGACAAGTGTGTGTGTGGGAGGGGACACTCCCACTGCAAGTGTGCAAGGGC 2300
QY 2371 CTCTACACCAAGCACCACCAACCAACCAAGTATTATCATGTTGTCACCATTCCTCTGGA 2430
Db 2301 ACCTTCTCGGCTCGCCCAAGAGCTTGGTTACATCAAGATGTTTGAATTCGCGCAGGA 2360
QY 2431 GCCCGGATTCGCTATGTAATGAAGCTCTCTACCTCTCTACATTTCTGT 2483
Db 2361 GCCAGACACCTGTAATTCAGGAAGCAGACACCAACCGGCATCACCTGGCCGT 2413

RESULT 12

US-09-369-364A-12
; Sequence 12, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369, 364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 5804
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(5648)
; NAME/KEY: misc_feature

; LOCATION: (1406)
; OTHER INFORMATION: n=T
; NAME/KEY: misc_feature
; LOCATION: (1563)
; OTHER INFORMATION: n=T
US-09-369-364A-12

Query Match 3.9%; Score 113.6; DB 4; Length 5804;
Best Local Similarity 47.7%; Pred. No. 3.8e-22;
Matches 542; Conservative 0; Mismatches 561; Indels 33; Gaps 6;

QY 1245 AGATCGGACTCGTCTATGACACCGCATCTTACTGACTGCTGGATATATGTTCTCGAA 1304
Db 974 AGTGGAAATCCATCATGATCTGCTGTTCTCTTAACAAGACAGGATATCTG---CAGAC 1030
QY 1305 GAATGAGCCCTGTGACACTTTTGGGATTTGCAACCCATAGTGAATGTGTAGTAATATCG 1364
Db 1031 TCACGACAAATGTGATACCTTAGGCTTGGCTGAACTGGGAACCAATTTGTGATCCCTATAG 1090
QY 1365 CAGCTGCGAGGATTAATGAAGATACAGGTCTTGGACTGGCCTTACCATTTGCCATGAGTC 1424
Db 1091 AAGCTGTTCTATTAGTGAAGATAGTGGATTGAGTACAGCTTTTACGATGCGCCATGAGCT 1150
QY 1425 TGGACACAACTTTGGCATGANTCATGTCGAGAGGGAACATGTGTAAAGTCCGAGGG 1484
Db 1151 GGGCCATGTGTTAAACATGCTCATGATGACACAAACAAATGTABAGAAGAGGATTAA 1210
QY 1485 CA-----ACATCATGTCCCTTACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTC 1535
Db 1211 GAGTCCCGCAGCATGTCTGGCTCCAACTTCTACACCAACCCCTGATGTGGTC 1270
QY 1536 ACCCTGCGAGCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTGC 1595
Db 1271 AAGTGTAGTCAAAATATATACCTGAGTTTTTACACACTGGTTATGGGAGTGTTCCT 1330
QY 1596 TGATCAGCCAAAGCCGTGTGAAGAAATACAAGTATCTTGAGAAATTCGCGAGGAAATATA 1655
Db 1331 TAAAGCACTGATC---CAGACCCTACCTTTGCTGTCCAACTGCCAGGATCCTTTA 1387
QY 1656 TGATGCAACACACAGTGCAGTGCAGTTCGAGAGAAAGCAAGCTCTGATGCTGGA 1715
Db 1388 CAACGTGAATAAACAATGNAATTTGATTTTGGACAGGTTCTCAGGTGTGCCATATAT 1447
QY 1716 CTTTAAAGGACATCTGTAAAGCCCTGTGGTGCCATCGTATTGGAAGAAATGTGAGAC 1775
Db 1448 GATGAGTGCACACCGCTCTGTGCAATTAACGTCAATGAGTACACAAAGGCTGCCGAC 1507
QY 1776 TAAATTTATGCCAGCAGCAGAGGACAAATTTTGGGCGATGACATGTGTGCGGGAGG 1835
Db 1508 TCAGCACACACCTGGGCGCATGGGACGAGTGCAGGCTGGAAAGCACTGCAAGNATGG 1567
QY 1836 ACAGTGTGTAATATGTTGATGAGGCGCCCAAGCCCAAGCTGCTCGGACTG 1895
Db 1568 ATTTGTGTTCCCAAGAAATGGATGTCCCGTG---ACAGATGGATCTCTGGGAGTTG 1624
QY 1896 GTCTTCTGTCCTCCATGCTCCAGGACCTCGGAGGGGAGTATCTCATAGAGTGCCT 1955
Db 1625 GAGTCCCTTTGGNACCTGCTCCAGAACATGTGGAGGGGSCATCAAAACAGCCATTCGAGA 1684
QY 1956 CTGCAACCAACCCCAAGCCATGCGATGGAGGAAAGTTCTGTGAGGCTCCACTCGCACTCT 2015
Db 1685 GTGCAACAGACCCAGAACCAAAATGTTGGAAAAATACTGTGTAGGACCTAGAAATTA 1744
QY 2016 GAAGCTCTGCAACAGTTCAGAAATGTCCCGGCGAGCTGTTGACCTCCGTGCTGCTCAGTG 2075
Db 1745 TAAGTCTGCAACCGGAGCCATGTCTCAAGCAGAGCGAGACTTCCGAGATGAACAGTG 1804
QY 2076 TGCCGAGCACAAACAGCAGACGAT-----TCAGAGGGCGGCACTACAAGTGGAAAGCT-- 2127
Db 1805 TGCTCACTTTGACGGGAAGCATTTTAAACATCAACGGTCTGCTTCCCATATGTGCGCTGGT 1864
QY 2128 -----TACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCCGAGAGG 2180

Qy	1310	AGCCCTGTGACACTTTGGGATTTGCACCCATAGTGGAAATGTTAGTAATAATTCGCAGCT	1369
Db	526	AGACATGTGATACCTCTTGGAGTGGCTGATGTTGGAACCTGTGTGTATCCGAGCAGAAGCT	585
Qy	1370	GCACGATTAATGAGATACAGGTCITGGACTGGCCTTCACCATTCGCCATGAGCTCGGAC	1429
Db	586	GCTCCGTCATAGAGATGATGGTTTACAAGCTGCCTTCACCACAGCCCCAATGAAATAGGCC	645
Qy	1430	ACAACCTTGGCATGATTATGATGGAGAAGGAAC-----ATGTGTAAAAAGT	1477
Db	646	ACGTGTTTAAACATGCCACATGATGCAAGACGATGTCGCCAGCCTTAATGTGTGAACC	705
Qy	1478	CCGAGGGCAACATCATGTCCCTACATTTGGCAGGACGCAATGGAGTCTTCTCTGTGTAC	1537
Db	706	AGGATTCACACATGATGGGCTCAATGCTTTCCAACTGGACACACAGCCATGGCTCTC	765
Qy	1538	CTGTGAGCCGCCAGTATCTACACAATTTCTAAGCACCGCTCAAGCTATCTGCCCTGTG	1597
Db	766	CTTGCAGTGCCTACATGATTACATATTTCTGGATAATGGTCATGGGGAATCTTTGATGG	825
Qy	1598	ATCAGCCAAAGCCTGTGAAGGAATACAAGTATCTCTGAGAAAATGCCAGGAGAATTATATG	1657
Db	826	A-----CAGGCTCAGAAATCCCATACAGTCTCCAGGGCGATCTCCCTGGCACCCTGTACG	879
Qy	1658	ATGCAAAACACACAGTSCAAGTGGCAGTTCCGGAGAGAAAGCCCAAGCTCTG---CATGCTGG	1714
Db	880	ATGCCAAACGGCAGTSCCAGTTTACATTTTGGCGAGGACTCCAAAACATCTCCCCGATGCG	939

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2003, 18:50:06 ; Search time 751 Seconds
(without alignments)
10337.968 Million cell updates/sec

Title: US-09-981-151A-7

Perfect score: 2895
Sequence: 1 cgcctctgatgaagcccg.....ccgggggtcttcctgcga 2895

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 134090451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/PCT NEW PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT NEW PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06 NEW PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07 NEW PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08 NEW PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09 NEW PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10 NEW PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60 NEW PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1874.6	64.8	3675	13	US-10-217-774-3
2	1874.6	64.8	3675	15	US-10-296-616-1
3	1874.6	64.8	4042	13	US-10-217-774-5
4	708	24.3	3666	12	US-10-240-545A-1
5	673.4	23.3	2589	14	US-10-226-560-1
6	673.4	23.3	3013	14	US-10-226-560-3
7	663	22.9	1476	13	US-10-217-774-1
8	292.8	10.1	525	12	US-10-029-386-8689
9	292.6	10.1	348	12	US-10-029-386-22417
10	267.4	9.2	339	10	US-09-833-381-1437
11	265.8	9.2	2052	14	US-10-103-377C-3
12	265.8	9.2	2377	14	US-10-103-377C-1
13	241.2	8.3	2217	14	US-10-226-560-6
14	227	7.8	2274	10	US-09-963-791-23
15	227	7.8	2727	10	US-09-963-791-1
16	227	7.8	2879	10	US-09-788-043C-2

17	225.4	7.8	1530	14	US-10-226-560-4	Sequence 4, Appli
18	215.4	7.4	5338	10	US-09-788-043C-6	Sequence 6, Appli
19	208	7.2	3132	9	US-09-858-081-10	Sequence 10, Appli
20	208	7.2	3329	9	US-09-858-081-8	Sequence 8, Appli
21	208	7.2	3776	12	US-09-842-469-1	Sequence 1, Appli
22	206.2	7.1	3207	9	US-09-858-081-3	Sequence 3, Appli
23	206.2	7.1	3402	9	US-09-858-068-3	Sequence 3, Appli
24	206.2	7.1	3403	9	US-09-858-081-1	Sequence 1, Appli
25	206.2	7.1	3403	9	US-09-858-068-1	Sequence 1, Appli
26	204.4	7.1	2445	9	US-09-858-081-13	Sequence 13, Appli
27	204.4	7.1	3471	9	US-09-858-081-11	Sequence 11, Appli
28	189.8	6.6	3414	12	US-09-842-469-3	Sequence 3, Appli
29	182.6	6.3	3218	10	US-09-918-171A-6	Sequence 6, Appli
30	181.4	6.3	3263	10	US-09-788-043C-4	Sequence 4, Appli
31	178	6.1	476	11	US-09-918-995-15335	Sequence 15335, A
32	176.8	6.1	3885	10	US-09-918-171A-16	Sequence 16, Appli
33	166.6	5.8	2190	14	US-10-163-547-17	Sequence 17, Appli
34	166.6	5.8	2925	14	US-10-163-547-15	Sequence 15, Appli
35	153.2	5.3	2270	12	US-10-188-869-5	Sequence 5, Appli
36	153.2	5.3	3369	12	US-10-188-869-11	Sequence 11, Appli
37	133.2	5.3	5004	12	US-10-188-869-9	Sequence 9, Appli
38	144	5.0	2848	10	US-09-918-171A-4	Sequence 4, Appli
39	138	4.8	2339	12	US-10-188-869-7	Sequence 7, Appli
40	138	4.8	3438	12	US-10-188-869-19	Sequence 19, Appli
41	123	4.2	3160	10	US-09-963-791-25	Sequence 25, Appli
42	119.2	4.1	1810	10	US-09-833-381-1056	Sequence 1056, Ap
43	118.4	4.1	6691	12	US-10-301-822-3	Sequence 3, Appli
44	118	4.1	1317	10	US-09-963-791-21	Sequence 21, Appli
45	118	4.1	1770	10	US-09-963-791-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-10-217-774-3
; Sequence 3, Application US/10217774
; Publication No. US20020193583A1

GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-217-774-3

Query Match	64.8%	Score 1874.6	DB 13	Length 3675
Best Local Similarity	84.4%	Pred. No. 0		
Matches 2287	Conservative	0	Mismatches 179	Indels 245
Gaps				
Qy	202	GCAGGCACCTGCTGGCGCCATGGGACCCGCGGAGCGGCTGGAGCCCGAGGTCCTCC	261	
Db	69	GCAGGCACCTGCTGGCGCCATGGGACCCGCGGAGCGGCTGGAGCCCGAGGTCCTCC	128	
Qy	262	GCCTCTCTCTCCACCCCGGAGCGGCTGGATGGAAGGAGGACATGATGGA	321	
Db	129	GCCTCTCTCTCCACCCCGGAGCGGCTGGATGGAAGGAGGACATGATGGA	175	
Qy	322	GCTGGAAACCATCGTCTCTAGCAAACTAACACAGGACAGAAAACCAAACTGATGTT	381	
Db	176	-----	175	

QY 382 CTCACATCAATGACCTGGTCTCTGCGCTACGAGGTGACCAACAGGGGCGATTACGTGCC 441
DB |||||
176 -----AATATGACCTGGTCTCTGCGCTACGAGGTGACCAACAGGGGCGATTACGTGCC 228
QY 442 CATGAATCATGACCATCAGCGCGGAGAGAGACAGTGGCGGTCCGAGGTGAGTCT 501
DB |||||
229 CATGAATCATGACCATCAGCGCGGAGAGAGAGTGGCGGTCCGAGGTGAGTCT 288
QY 502 CTTCACTCTCGGCTGAAGGCCGCCAGGACGAGCTTCCATGGATCTGAGGACTTCCAGC 561
DB |||||
289 CTTCACTCTCGGCTGAAGGCCGCCAGGACGAGCTTCCATGGATCTGAGGACTTCCAGC 348
QY 562 AGCTAGTGGTCTGCTGCTTTATTTGTCGAGAGCTTGGAAAGACAGGACTAAGTCTGTG 621
DB |||||
349 AGCTAGTGGTCTGCTGCTTTATTTGTCGAGAGCTTGGAAAGACAGGACTAAGTCTGTG 408
QY 622 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGCATCACACAGAAAC 681
DB |||||
409 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGCATCACACAGAAAC 468
QY 682 TCGGCATCGCATGAGGAGAGTTCTGTGAGGCTTCACTCGCATCTGAAGCTCTGCAAC 741
DB |||||
469 TCCTCA-----GTGCGCCCTTTCAACCTGCCAAGGCTTGTC 504
QY 742 AGTCAGAAATGTCGCCGAGACAGTGTGACTTCCGTGCTGCTCAGTGTGCGAGCACAC 801
DB |||||
505 GGCATGATCGAACAGAGAGGAGATTACTTCTTAAGGCGACTTCTCTCACACCTTCA 564
QY 802 AGCAGACGATTCAGAGCGCGGACACTACAAGTGG--AAGCCCTTACACTCAAGTGAAGCCG 859
DB |||||
565 TGGAACTCGGACAGCTGCCAAGGAGCTCGCCATCCACGCTACTGTACAGAGATCC 624
QY 860 ACTTATGCAACTTACTGTATCGCAGAGATTTGATTTCTTTCTTTTGTGCAATA 919
DB |||||
625 ACAGAGCCCATGCTCTCTGGGGCAGTGAGGTCTCTGCTGACCTC-----AAGG 672
QY 920 AAGTCAAAAGATGGGACTCCATGCTCGGAGGATAGCGTAAATGTTGTATAGATGGGAT 979
DB |||||
673 ACATGGAGCTGGACATCAACCCCTGCAAGAGGAGACCTTGCCTGGGACTGCCACAA 732
QY 980 GTGAGCTCATGTGGTGTCCACATCTGC-GCACATGCCCGAGCTCCCAAGGAGACCTC 1038
DB |||||
733 AAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCAGCTCCCAAGGAGACCTC 792
QY 1039 TTCACTTTGCCAGATGATAGTCTTGTCTTACGGCATAGGCGCTCTCTCTGAGGTCC 1098
DB |||||
793 TTCACTTTGCCAGATGATAGTCTTGTCTTACGGCATAGGCGCTCTCTCTGAGGTCC 852
QY 1099 CATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGCACAAAGAAAGATGATGCAA 1158
DB |||||
853 CATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGCACAAAGAAAGATGATGCAA 912
QY 1159 AAGCAGTGGCATGAAGAAATACCACTACCTGCTACGATCAAGTATCAACATGGTATCTGT 1218
DB |||||
913 AAGCAGTGGCATGAAGAAATACCACTACCTGCTACGATCAACATGGTATCTGT 972
QY 1219 TTATTTCAAAGA----- 1229
DB |||||
973 TTATTTCAAAGATGMAAATAAGGAGGAAACATCAACTTCGAATTTAGGTCTGATCTT 1032
QY 1230 ----- 1229
DB 1033 CTAGAAATGAACAGCGAGGACTGGTATAGTCAACCAAGCAGACACACCTTAAGTAGC 1092
QY 1230 -----TGATTTGATGGGAAGATGGGACTCGTCAATGACCAAGCCATC 1272
DB |||||
1093 TTCTGCAGTGGCAGTCTGATTTGATGGGAAGATGGGACTCGTCAATGACCAAGCCATC 1152
QY 1273 TTACTGACTGGTCTGGATATATGTTCTGGAAGAAATGAGCCCTGTGACACTTTGGGATTT 1332
DB |||||
1153 TTACTGACTGGTCTGGATATATGTTCTGGAAGAAATGAGCCCTGTGACACTTTGGGATTT 1212

QY 1333 GCACCCATTAAGTGGAAATGTAGTAAATATFCGAGCTGCACGANTTAATGAAGATACAGGT 1392
DB |||||
1213 GCACCCATTAAGTGGAAATGTAGTAAATATFCGAGCTGCACGANTTAATGAAGATACAGGT 1272
QY 1393 CTTGGACTGGCTTCCACTTCCCATGAGTCTGGACACACCTTTGGCATGATTCATGAT 1452
DB |||||
1273 CTTGGACTGGCTTCCACTTCCCATGAGTCTGGACACACCTTTGGCATGATTCATGAT 1332
QY 1453 GGAGAAAGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCTCATTCATGGCAGGA 1512
DB |||||
1333 GGAGAAAGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCTCATTCATGGCAGGA 1392
QY 1513 CGCAATGGAGTCTTCTCTGGTCAACCTCGAGCGCCAGTATCTACACAAATTTCTAAGC 1572
DB |||||
1393 CGCAATGGAGTCTTCTCTGGTCAACCTCGAGCGCCAGTATCTACACAAATTTCTAAGC 1452
QY 1573 ACCGCTCAAGCTATCTGCTTGCCTGATCAGCCAAAGCCCTGTGAAGGAATACAAAGTATCT 1632
DB |||||
1453 ACCGCTCAAGCTATCTGCTTGCCTGATCAGCCAAAGCCCTGTGAAGGAATACAAAGTATCT 1512
QY 1633 GAGAAATTCGACGAGAGATTTATATGATGCAAAACACACAGTGCAGGTGGCAGTTCCGAGAG 1692
DB |||||
1513 GAGAAATTCGACGAGAGATTTATATGATGCAAAACACACAGTGCAGGTGGCAGTTCCGAGAG 1572
QY 1693 AAGCCCAAGCTCTGCTGATGTGAGACTTTAAAGGAGACATCTGTAAAGCCCTGTGGTCCCAT 1752
DB |||||
1573 AAGCCCAAGCTCTGCTGATGTGAGACTTTAAAGGAGACATCTGTAAAGCCCTGTGGTCCCAT 1632
QY 1753 CGTATTGGAAGAAATGTGAGACTTAAATTTATGTCAGCAGCAGAGAGGCAAAATTTGTGG 1812
DB |||||
1633 CGTATTGGAAGAAATGTGAGACTTAAATTTATGTCAGCAGCAGAGAGGCAAAATTTGTGG 1692
QY 1813 CATGACATGTGTGTCGCGGAGGACAGTGTGAAATATGTTGATGAAGGCCCCCAAGCCC 1872
DB |||||
1693 CATGACATGTGTGTCGCGGAGGACAGTGTGAAATATGTTGATGAAGGCCCCCAAGCCC 1752
QY 1873 ACCCATGGCCACTGTGTGGACTTGTCTTGGTCCCATGTCTCAGGACCTGTGGAGGG 1932
DB |||||
1753 ACCCATGGCCACTGTGTGGACTTGTCTTGGTCCCATGTCTCAGGACCTGTGGAGGG 1812
QY 1933 GGAGTATCTCATAGGAGTCCCTCTGCACCAACCCCAAGCCCATCGATGAGGAGGAGTTTC 1992
DB |||||
1813 GGAGTATCTCATAGGAGTCCCTCTGCACCAACCCCAAGCCCATCGATGAGGAGGAGTTTC 1872
QY 1993 TGTGAGGCTCCACTCGCATCTGAAGCTCTCAACAGTCTCAAAATGTCCCGGGACACT 2052
DB |||||
1873 TGTGAGGCTCCACTCGCATCTGAAGCTCTCAACAGTCTCAAAATGTCCCGGGACACT 1932
QY 2053 GTTGACTTCCGCTGCTCAGTGTCCGAGCAGACACAGACAGGATTCAGAGGGCGGCAC 2112
DB |||||
1933 GTTGACTTCCGCTGCTCAGTGTCCGAGCAGACACAGACAGGATTCAGAGGGCGGCAC 1992
QY 2113 TACAAGTGGAAAGCTTACACTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 2172
DB |||||
1993 TACAAGTGGAAAGCTTACACTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 2052
QY 2173 GCAGAGGATTTGATTTCTTTTCTTTTGTGCAATTAAGTCAAAAGATGGGACTCCATGC 2232
DB |||||
2053 GCAGAGGATTTGATTTCTTTCTTTTGTGCAATTAAGTCAAAAGATGGGACTCCATGC 2112
QY 2233 TCGGAGGATAGCCGTAATGTTTGTATAGATGGGATGTGAGAGGTTGGATGTGACAT 2292
DB |||||
2113 TCGGAGGATAGCCGTAATGTTTGTATAGATGGGATGTGAGAGGTTGGATGTGACAT 2172
QY 2293 GTCTTTGGATCTGATGTGTTGAAGACCTCTGTGGGGTGTGTAAACGGGAATACTCAGCC 2352
DB |||||
2173 GTCTTTGGATCTGATGTGTTGAAGACCTCTGTGGGGTGTGTAAACGGGAATACTCAGCC 2232
QY 2353 TGCAGATTCACAGGGCTCTTACACCAAGCAGCACCACCAACCAAGTATTTACATG 2412
DB |||||
2233 TGCAGATTCACAGGGCTCTTACACCAAGCAGCACCACCAACCAAGTATTTACATG 2292
QY 2413 GTCACCATTCCTTCTGGAGCCCGGAGTATCGCATCTATGAATGAACGCTCTTACCTCC 2472

Db 2293 GTACACATTCCTTCGTGGAGCCCGGAGATCCCGCATCTATGAATGAACGCTCTACCTCC 2352
 Qy 2473 TACATTTCTGTGGCAATGCCCTCAGAAAGTACTACCTGAATGGCACTGGACCGTGGAC 2532
 Db 2353 TACATTTCTGTGGCAATGCCCTCAGAAAGTACTACCTGAATGGCACTGGACCGTGGAC 2412
 Qy 2533 TGGCCCGCCGGTACAAATTTTCGGGCACACTCTTCGACTACAGACGGTCTCTAATGAG 2592
 Db 2413 TGGCCCGCCGGTACAAATTTTCGGGCACACTCTTCGACTACAGACGGTCTCTAATGAG 2472
 Qy 2593 CCGAGAACTTAATCGCTACTGAGCAACCAACAGAGACACTGATTTGGAGCTGCTGTTT 2652
 Db 2473 CCGAGAACTTAATCGCTACTGAGCAACCAACAGAGACACTGATTTGGAGCTGCTGTTT 2532
 Qy 2653 CAGGGAAGGAGCCCGGCTTGGCTGGGAATACCTGATGCTGCTTGGGACCGAGAG 2712
 Db 2533 CAGGGAAGGAGCCCGGCTTGGCTGGGAATACCTGATGCTGCTTGGGACCGAGAG 2592
 Qy 2713 CAGCCCGCTGCCAGCCAGCTACACTTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCT 2772
 Db 2593 CAGCCCGCTGCCAGCCAGCTACACTTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCT 2652
 Qy 2773 TGGGAGGGGG 2783
 Db 2653 TGGGAGGGGG 2663

RESULT 2

US-10-296-616-1
 ; Sequence 1, Application US/10296616
 ; Publication No. US20030129658A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
 ; TITLE OF INVENTION: No. US20030129658A1el1 protease
 ; FILE REFERENCE: Y0132PCT-664
 ; CURRENT APPLICATION NUMBER: US/10/296,616
 ; CURRENT FILING DATE: 2002-11-26
 ; PRIOR APPLICATION NUMBER: JP 2000-393372
 ; PRIOR FILING DATE: 2000-12-25
 ; NUMBER OF SEQ ID NOS: 26
 ; SEQ ID NO 1
 ; LENGTH: 3675
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(3675)
 US-10-296-616-1

Query Match 64.8%; Score 1874.6; DB 15; Length 3675;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;
 Qy 202 GCAGGACCTCGCTGGCCCATGGGACCCGAGCGGAGCGCTGGGAGCCCGAGCGTCCC 261
 Db 69 GCAGGACCTCGCTGGCCCATGGGACCCGAGCGGAGCGCTGGGAGCCCGAGCGTCCC 128
 Qy 262 GCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAAAGGGCGGACATGATGAA 321
 Db 129 GCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAAAGGGCGG 175
 Qy 322 GCTGGAACCATCTGTTCTCAGCAACTACACAGGAACAGAAACCAACACTGCATGTT 381
 Db 176 ----- 175
 Qy 382 CTCACATATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
 Db 176 -----AATATGACCTGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 228
 Qy 442 CATGAATCATGACCATGACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
 Db 229 CATGAATCATGACCATGACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288

Qy 502 CTTACCTTCGGCTGAAAGGCCCGGACGACGACTTCCACATGATCTGAGGACTTCCAGC 561
 Db 289 CTTACCTTCGGCTGAAAGGCCCGGACGACGACTTCCACATGATCTGAGGACTTCCAGC 348
 Qy 562 AGCCTAGTGGCTCCTGGCTTTATTTGTCAGAGCTTTGGGAAAGACAGGACCTAAGTCTGTG 621
 Db 349 AGCCTAGTGGCTCCTGGCTTTATTTGTCAGAGCTTTGGGAAAGACAGGACCTAAGTCTGTG 408
 Qy 622 CAGACTTTACCGCCAGAGGACTTCTGTTCTATCAAGGCTTTTGGGATCAACAGAAAC 681
 Db 409 CAGACTTTACCGCCAGAGGACTTCTGTTCTATCAAGGCTTTTGGGATCAACAGAAAC 468
 Qy 682 TCGCCATCGCATGGAGGAGTTCTGTAGGGCTCCACTCGCACTCTGAAGCTCTTCAAC 741
 Db 469 TCTCTCA -----GTGGCCCTTTCAACCTGCCAAGGCTTGTCA 504
 Qy 742 AGTCAGAAATGTCCCGGACAGCTGTGACTTCCGTCGTGCTGCTGCTGCTGCTGCTGCTGCT 801
 Db 505 GGCATGATACGACAGAGAGAGGACGATTTACTTCTTAAGGCCACTTCTCTTCAACCTCTCA 564
 Qy 802 AGCAGAGATTCAGAGGCGGACACTACAGTGG--AAGCCTTACACTCAAGTAGAGCCG 859
 Db 565 TGGAACTCGGACAGGCTGCCAAGGACGCTGCCATCCACGCTACTGTACAGAGANTCC 624
 Qy 860 ACTTATGCAAACTCTACTGTATCGCAGAAAGGATTGATTTCTTCTTTTCTTTTCTCAATA 919
 Db 625 ACAGAGCCCATGCTCTCGGGCCACTGAGGCTCTGTGACCTC-----AAGG 672
 Qy 920 AAGTCAAAGATGGGACTCTCATGCTCGAGAGATAGCGGTAAATGTTGTATAGATGGGATAT 979
 Db 673 ACATGGGAGCTGGCAGATCAACCCCTGCACAGCAGCGACCTTCGCCCTGGGACTGCCACAA 732
 Qy 980 GTGAGCTCAGTGTGTGTCACATCTGC--GCACATGCCCGAGCTCCCAAGGAGACCTC 1038
 Db 733 AAGCAGATTTCTGTGAGAGCCAGAAATACATGCCCGAGCTCCCAAGGAGACCTC 792
 Qy 1039 TTCACTTTGCCAGATGATATAAGTCTTGTCTTACGGCATAAGCGCTCTCTTCTGAGGTCC 1098
 Db 793 TTCACTTTGCCAGATGATATAAGTCTTGTCTTACGGCATAAGCGCTCTCTTCTGAGGTCC 852
 Qy 1099 CATAGAATGAAGNACTGAACGTTGGAGAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1158
 Db 853 CATAGAATGAAGNACTGAACGTTGGAGAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 912
 Qy 1159 AACCATGGCCATGAAAATATCACCACTACGTCCTACGATCTCAACATGGTATCTGCT 1218
 Db 913 AACCATGGCCATGAAAATATCACCACTACGTCCTACGATCTCAACATGGTATCTGCT 972
 Qy 1219 TTATTCAAGA----- 1229
 Db 973 TTATTCAAGAATGGAACAATAGGAGGAAACATCAACATTTGCAATTTAGGCTCTGATCTT 1032
 Qy 1230 ----- 1229
 Db 1033 CTAGAAGATGAACAGCCAGGACTGGTGATTAAGTCAACCGAGACCCACACCTTAAGTAGC 1092
 Qy 1230 -----TGAATTGATGGGAAAGATGGGACTCGTCAATGACCAACCGCATC 1272
 Db 1093 TTCTGCGAGTGGCAGTCTGGATTTGATGGGAAAGATGGGACTCGTCAATGACCAACCGCATC 1152
 Qy 1273 TTACTGAGTGGTCTGGATATATGCTCTGGAAGATGAGCCCTGTGACACTTTGGGATTT 1332
 Db 1153 TTACTGAGTGGTCTGGATATATGCTCTGGAAGATGAGCCCTGTGACACTTTGGGATTT 1212
 Qy 1333 GCACCCATTAAGTGGAAATGTAGTAATATCGCAGCTGCACGATTAATGAAGATACAGGT 1392
 Db 1213 GCACCCATAAGTGGAAATGTAGTAATATCGCAGCTGCACGATTAATGAAGATACAGGT 1272
 Qy 1393 CTTGGACTGCGCTTCCACCATGGCCATGAGTCTGGACACAACTTTGGGATGATCTGAT 1452
 Db 1273 CTTGGACTGCGCTTCCACCATGGCCATGAGTCTGGACACAACTTTGGGATGATCTGAT 1332

Db 447 AGCTACTGGCTCTGCTCTTTATTTGTCAGACGCTGGGAAAGACAGCGACTAAGTCTGTG 506
Qy 622 CAGACTTTACCGCCAGAGGACTCTGTTTCTATCAAGGCTCTTTGGGATCACAAGAAAC 681
Db 507 CAGACTTTACCGCCAGAGGACTCTGTTTCTATCAAGGCTCTTTGGGATCACAAGAAAC 566
Qy 682 TCGCCATCGCATGAGGAGAAAGTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAAC 741
Db 567 TCCTCA-----GTGGCCCTTTTCAACCTGCCAAGGCTTGTCA 602
Qy 742 AGTCAGAAAGTCCCGGGACAGTGTGACTCTCGTGTCTCAGTGTGCGGACACAC 801
Db 603 GGCATGATACGAACAGAGAGGCGAGATTTCTTAGGCCACTTCTCTCACCCTCTCA 662
Qy 802 AGCAGACGATTCAGAGGGCGGCACTACAAGTGG--AAGCCTTACATCAAGTAGAAGCGG 859
Db 663 TGGAAACTCGGCAGAGCTGCCAAGGAGCTCGCCATCCACGTACTGTACAGAGATCC 722
Qy 860 ACTTATGCAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTCTTGTCAATA 919
Db 723 ACAGAGCCCATGCTCTGGGGCCAGTGAGGCTCTGTGTACCTC-----AAGG 770
Qy 920 AAGTCAAGATGGGACTCCATGCTCGAGGATACCGTAATGTTGTATAGATGGATAT 979
Db 771 ACATGGGAGCTGGCACATCAACCTCTGCACAGAGCGACCTTCGCTGGGACTGCCACA 830
Qy 980 GTGAGCTCAGTGTGTGTCACATCTGC-GCACATGCCCGAGCTCCAGGAGAGCCTC 1038
Db 831 AAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCTCCAGGCTCCAGGAGAGCCTC 890
Qy 1039 TTCATCTGGCAGATGAGTATAAGTCTTGTCTTACGGCATAAAGGCTCTCTTTCTGAGTCC 1098
Db 891 TTCATCTGCCAGATGAGTATAGTCTTGTCTACGGCATTAAGGCTCTCTTTCTGAGGTC 950
Qy 1099 CATGAAATGAAGAACTGAACGCTGGAGACCTTTGGTGTGTGTCGACAAAGATGTCGAA 1158
Db 951 CATGAAATGAAGAACTGAACGCTGGAGACCTTTGGTGTGTGTCGACAAAGATGTCGAA 1010
Qy 1159 AACCATGGCCATGAATATACCACTTACCTGCTCAGCATACTCAACATGTTGATCTGCT 1218
Db 1011 AACCATGGCCATGAATATACCACTTACCTGCTCAGCATACTCAACATGTTGATCTGCT 1070
Qy 1219 TTATTCAAAGA----- 1229
Db 1071 TTATTCAAAGATGGAACAATAGGAGAAACATCAACATTCGAATTTGAGTCTGTATCTT 1130
Qy 1230----- 1229
Db 1131 CTAGAGATGAACAGCCAGGACTGTTGATAGTCAACCCGACGACACACCTTAAGTAGC 1190
Qy 1230-----TGATTTGATGGGAAAGATGGGACTCGTCTATGACCCAGCCATC 1272
Db 1191 TTCTGCCAGTGGCAGTCTGGATTTGATGGGAAAGATGGGACTCGTCTATGACCCAGCCATC 1250
Qy 1273 TTACTGACTGTCTGGATATGTTCTTGGGAAGATGAGCCCTCTGTGACACTTTGGGATTT 1332
Db 1251 TTACTGACTGTCTGGATATGTTCTTGGGAAGATGAGCCCTGTGACACTTTGGGATTT 1310
Qy 1333 GCACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGCAAGTAAATGAAGATACAGGT 1392
Db 1311 GCACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGCAAGTAAATGAAGATACAGGT 1370
Qy 1393 CTTGACCTGGCTTCCACATTTGCCATGAGTCTCGACACACTTTGGCATGATTCATGAT 1452
Db 1371 CTTGACCTGGCTTCCACATTTGCCATGAGTCTGGACACAACTTTGGCATGATTCATGAT 1430
Qy 1453 GGAGAGGGAACATGTGTAAGAGTCCGAGGCGCAACATCATGTCCTTACATTTGGCAGGA 1512
Db 1431 GGAGAGGGAACATGTGTAAGAGTCCGAGGCGCAACATCATGTCCTTACATTTGGCAGGA 1490
Qy 1513 CGCAATGAGATCTTCTCTGTCACCTTGGAGCGCCGCGATCTACACAAATTTCTAAGC 1572
Db 1571 CGCGAGAACTTAATCGCTACTGCGACCAACCAAGAGACACTGATTTGGAGCTGCTGTTT 2630

1491 GCGAATGGAGTCTTCTCTGTCACCTCGACGCGCGCAGTATCTACACAAATTTCTAAGC 1550
1573 ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGGAATACAAAGTATCCT 1632
1551 ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGGAATACAAAGTATCCT 1610
1633 GAGAAATGCCAGAGAAATATATGATGCAAAACACACAGTGAAGTGGAGTTCGAGAG 1692
1611 GAGAAATGCCAGAGAAATATATGATGCAAAACACACAGTGAAGTGGAGTTCGAGAG 1670
1693 AAAGCAAGCTCTCATGCTGGACTTTAAAGAGACATCTGTAAAGCCCTGTGTGCTCAT 1752
1671 AAAGCAAGCTCTCATGCTGGACTTTAAAGAGACATCTGTAAAGCCCTGTGTGCTCAT 1730
1753 CGTATTTGAAGAAATGTGAGACTAAATTTATGCGCAGCAGACAGAGCCAAATTTGTGG 1812
1731 CGTATTTGAAGAAATGTGAGACTAAATTTATGCGCAGCAGACAGAGCCAAATTTGTGG 1790
1813 CATGACATGTGCTCCGCGGAGGACAGTGTGTGAATATGCTGATGAAGCCCAAGGCC 1872
1791 CATGACATGTGCTCCGCGGAGGAGACAGTGTGTGAATATGCTGATGAAGCCCAAGGCC 1850
1873 ACCATGCGCACTGCTCGGACTGCTTCTGTCGCTCCGATGCTCCAGGACTCGCGAGGG 1932
1851 ACCATGCGCACTGCTCGGACTGCTTCTTGTGCTCCGATGCTCCAGGACTCGCGAGGG 1910
1933 GAGGTATCTCATAGAGTGGCTCTGCACCAACCCCAAGCCATCGCATGAGGAGGATTC 1992
1911 GAGGTATCTCATAGAGTGGCTCTGCACCAACCCCAAGCCATCGCATGAGGAGGATTC 1870
1993 TGTGAGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCTCCGAGCAGT 2052
1971 TGTGAGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCTCCGAGCAGT 2030
2053 GTTGACTTCCGCTGCTCAGTGTGCGAGGACAAACAGCAGCAGTTCAGAGGCGGAC 2112
2031 GTTGACTTCCGCTGCTCAGTGTGCGAGGACAAACAGCAGCAGTTCAGAGGCGGAC 2090
2113 TACAGTGGAAAGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATC 2172
2091 TACAGTGGAAAGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATC 2150
2173 GCAGAGATTTGATTTCTTCTTTTGTCAATTAAGTCAAGATGAGGACTCCTATGC 2232
2151 GCAGAGATTTGATTTCTTCTTTTGTCAATTAAGTCAAGATGAGGACTCCTATGC 2210
2233 TCGAGAGATAGCCGTAATGTTGTATAGATGGGATATGAGAGAGTGGATGTGACAAT 2292
2211 TCGAGAGATAGCCGTAATGTTGTATAGATGGGATATGAGAGAGTGGATGTGACAAT 2270
2293 GTTCTTGGATCTGATGCTGTTGAAGAGCTCTGTGGGTGTGTAACGGGAATTAACAGCC 2352
2271 GTTCTTGGATCTGATGCTGTTGAAGAGCTCTGTGGGTGTGTAACGGGAATTAACAGCC 2330
2353 TGACAGATTCAGAGGGTCTCTACCAAGCACCACACACCAACCAAGTATTATCATG 2412
2331 TGACAGATTCAGAGGGTCTCTACCAAGCACCACACACCAACCAAGTATTATCATG 2390
2413 GTTCAATTTCTTGGAGCCCGGAGTATCGCATCTATGAATGAAGCTCTTACTGCTC 2472
2391 GTTCAATTTCTTGGAGCCCGGAGTATCGCATCTATGAATGAAGCTCTTACTGCTC 2450
2473 TACATTTCTGTCGCAATGCTCAGAGGTACTTACTGATGCGCACTGGAGCGGTGGAC 2532
2451 TACATTTCTGTCGCAATGCTCAGAGGTACTTACTGATGCGCACTGGAGCGGTGGAC 2510
2533 TGGCCCGGCGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGAGCGGCTCTATATGAG 2592
2511 TGGCCCGGCGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGAGCGGTCTCTATATGAG 2570
2593 CCGGAGAACTTAATCGCTACTGAGCCAAACCAAGAGACACTGATTTGGAGCTGCTGCTTT 2652
2571 CCGGAGAACTTAATCGCTACTGAGCCAAACCAAGAGACACTGATTTGGAGCTGCTGCTTT 2630

QY 2653 CAGGAGGAAACCCGGGTGTTCCCTGGGATACCTCCATCCCTCGCTTGGGACCCAGAG 2712
Db 2631 CAGGGAAGAAACCCGGGTGTTCCCTGGGATACCTCCATCCCTCGCTTGGGACCCAGAG 2690
QY 2713 CAGCCCTTGGCCAGCCAGCTACACTTGGGACATCGTGGCTCTGAGTGTCTCCGTGTC 2772
Db 2691 CAGCCCTTGGCCAGCCAGCTACACTTGGGACATCGTGGCTCTGAGTGTCTCCGTGTC 2750
QY 2773 TCGGAGGGGG 2783
Db 2751 TCGGAGGGGG 2761

RESULT 4

US-10-240-545A-1
; Sequence 1, Application US/10240545A
; Publication No. US20030185828A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: No. US20030185828A1el agg:recanase
; FILE REFERENCE: 08959.0002
; CURRENT APPLICATION NUMBER: US/10/240,545A
; PRIOR FILING DATE: 2002-10-02
; CURRENT APPLICATION NUMBER: PCT/JP01/11033
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: JP 2000-384300
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 1
; LENGTH: 3666
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3666)
US-10-240-545A-1

Query Match 24.5%; Score 708; DB 12; Length 3666;

Best Local Similarity 64.9%; Pred. No. 2,4e-213;

Matches 1080; Conservative 0; Mismatches 580; Indels 4; Gaps 2;

QY 1124 AGACCTTGGTGGTGGCGAAGAAAGATGATGCAAAACCATGCGCATGAAATATCACCA 1183
Db 1014 AAACGTGGTGGTGGCGAAGAAAGATGATGCAAAACCATGCGCATGAAATATCACCA 1073
QY 1184 CCTAGCTGCTCAGCATACCTCAACATGGTATCTGCTTTTATCAAGATGATGATGGGA 1243
Db 1074 CCATCATGAGACCACTCTCTGAATAGTTTTTG-TCAATGGCAGTCTGCCCTCATGGAA 1132
QY 1244 AAGATGGGACTCGTCATGACCCGCGCATCTTACTGACTGGTCTGGATATATGTTCTTGA 1303
Db 1133 AGAATGGCAAGACATGATCATGCTTACTTAAACAGGATTTGATATTTGTTCTTGA 1192
QY 1304 AGAATGGCCCTGTCACACTTTGGGATTTGACCCATAGTGAATGATGATGATGATATC 1363
Db 1193 AGAATGAACCATGTGACACTCTAGGGTTTGCCCATCAGTGAATGTGCTCTAAGTACC 1252
QY 1364 GCAGCTGCACGATTAATGAAGATACAGGTCTTTGGACTGGCTTCCACCATGCGCCATGAGT 1423
Db 1253 GAAGTTGATCATCAATGAGGACACAGGACTTGGCTTGGCTTCCCTCATGCTCATGAGT 1312
QY 1424 CTGGACACAACTTTGGCATGATTCATGATGGAAGAGGAAACATGTGTAAAAGTCCGAGG 1483
Db 1313 CAGGACCAAACTTTGGTATGATTCACGACGGAAGAGGAAATCCCTGCAAGAAAGGCTGAAG 1372
QY 1484 GCAACATCATGTCCTCATATGCGGACGAGGACGCAATGAGTCTTCTCTGGTCCACCTGCA 1543
Db 1373 GCAATATCATGTCCTCCACACTGACCGGAAACAAATGAGGTGTTTTTCATGGTCTTCCCTGCA 1432
QY 1544 GCCGCGAGTATTCACAAATTTCTAAGCACCGCTCAAGCTATCTGCTCTGCTGATCAGC 1603
Db 1433 GCCGCGAGTATTCACAAATTTCTAAGCACCGCTCAAGCTATCTGCTCTGCTGATCAGC 1492

QY 1604 CAAAGCCTGTGACGAATACAAAGTATCTCTGAGAAATTTCCAGGAGAAATATATATGACAA 1663
Db 1493 CCAAGCAGCAGGACAGTATAAATATATCCGCAAACTACCAAGGACAGATTTATGATGCTG 1552
QY 1664 ACACACAGTGCAGTGGCAGTTCCGAGAGAAAGCAAGCTCTGATGCTGACCTTTAAAA 1723
Db 1553 ACACACAGTGCAGTGGCAGTTCCGAGAGAAAGCAAGTATGACGCTTGGTTTGTGA 1612
QY 1724 AGGACATCTGTPAAAGCCCTGTGTGCTCATCTATTTGGAAGAAATGTGAGACTAAATTA 1783
Db 1613 AGGATATTTGCAATCACTTTGTGTGCCACCGAGTAGGCCACAGGTGTGAGACCAAGTTA 1672
QY 1784 TCCAGCAGCAGAGAGGACACATTTGTGGGATGACATGTGTGTCGGGAGGAGCAGTGTG 1843
Db 1673 TGCCCGCAGCAGAGAGGAGACCTTTGTGTGATGATGTGTGTGTCGCGCAGGCTGCG 1732
QY 1844 TGAATATGTGATGAAGGCCCCCAAGCCCACTGCGCACTGCTGCGACTGTGCTCTT 1903
Db 1733 TAAAGTTTGGGAGCTCGGCCCCCGGCCCATCCACGCCAGTGTGCTCGCTGCTGAAGT 1792
QY 1904 GGTCCCATGCTCCAGGACCTGCGGAGGGGAGTATCTCATAGAGTGTGCTCTGCAACCA 1963
Db 1793 GGTGAGATGTTCGCGACATGTGTGAGGAGTCAAGTTCCAGGAGAGACACTGCAATA 1852
QY 1964 ACCCAAGCCATCGCATGGGAGGAGTTCTGTGAGGCTCCACTCGGACTCTGAGCTCT 2023
Db 1853 ACCCAAGCCCTCAGTATGTGTGCTTATCTGTCCAGGTCTGAGCCGTATTTATCAGCTGT 1912
QY 2024 GCAACAGTCAAGAAATGTCCCGGGACAGTGTGTGACTTCCGCTGTGCTCAGTGTGCCGAGC 2083
Db 1913 GCAATATTAACCTTGAATGAATAATAGTTGGATTTTCGGGCTCAACAGTGTGCAAGT 1972
QY 2084 ACAACAGCAGAGATTCAGAGGGGCGCATCTAAGTGTGAGGCTTACACTCAAGTAGAG 2143
Db 1973 ATAAACAGCAACCTTTCCGTGGATGTTTACCAAGTGGAAACCCCTATACAAAAGTGAAG 2032
QY 2144 ATCAGGACTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTCTTTCTTGT 2203
Db 2033 AGGAAGATCGATGCAAACTGTACTGTCAAGCTGAGACTTGAATTTTTTTTGTGCAATGT 2092
QY 2204 CAATAAAGTCAAAAGATGGGACTCCATGCTGGAGGATAGCCGTAATTTGTATAGATG 2263
Db 2093 CCGCAAGTGAAGATGGAACCTCCCTGCTCCCCCAACAAATATGATTTGTATTGACG 2152
QY 2264 GGATATGTGAGAGTGTGATGTGACAAATGCTCTGATGCTGTGATGCTGTGAGACGCT 2323
Db 2153 GGGTTGTGAACATAGTGGGATGTGATCATGAACCTAGGCTCTAAAGCAGTTTTCAGATGCTT 2212
QY 2324 GTGGGTGTGTAAAGGCAATAAAGTCAAGCTGACGATTCAGGGGTCTCTACACCAAGC 2383
Db 2213 GTGGGTGTGCAAAAGTGAATAATCACTTGAAGTTTAAAGGCTGTGATCTCAAC 2272
QY 2384 ACCACACACCAACAGATTTATCAATGTCACCTTCTCTGAGGAGGAGGAGTATCC 2443
Db 2273 AGCATAAAGCAAAATGAAATATATCCGGTGTGTCATCTCCAGCTGGCGCGGAGATCG 2332
QY 2444 GCATCTATGAAATGAACGCTCTACCTCTACATTTCTGTGCGCAATGCGCTCAGAGGT 2503
Db 2333 AAATCCAGGAGCTGACGATTTCTTCCAGTTTACCTTGCAGTTCGAGGCTCAGTCAAAAGT 2392
QY 2504 ACTACCTGAATGGGCACTGGAACCTGTGAGCTGGCGCGCGCGGTACAAATTTTCGGGCACTA 2563
Db 2393 ATTACCTCACCGGGGCTGGAGCATCGACTGGCTTGGGAGTTTCCCTTCCCTGCTGGACCA 2452
QY 2564 CTTTCCGACTACAGACGCTCTCTATTAATGAGCCCGAGAACTTAATGCTGTACTGAGACCA 2623
Db 2453 CGTTGTAATACAGCGCTCTTTCAACCGCGCGGAACTGTGTACGCGCCAGGCGCCACAA 2512
QY 2624 ACAGACACTGATTTGTGAGCTGCTGTTTCCAGGAGGAGAACCCGGGTGTGCTGCTGGAGT 2683
Db 2513 ATGAGCGCTGTGCTTTGAAATTTCTGATGCAAGGCAAAATCCAGGAGTGTGCTGAGAGT 2572

US-10-226-560-3
; Sequence 3, Application US/10226560
; Publication No. US20030050464A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Nepomniichy, Boris
; APPLICANT: Walke, D. Wade
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20030050464A1el Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0372-USA
; CURRENT APPLICATION NUMBER: US/10/226,560
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US 60/314,049
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 09/917,614
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/221,644
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3013
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-226-560-3

Query Match 23.3%; Score 673.4; DB 14; Length 3013;
Best Local Similarity 65.6%; Pred. No. 2.1e-202;
Matches 997; Conservative 0; Mismatches 521; Indels 1; Gaps 1;
QY 1124 AGACCTTGGTGTGTCGACAAAGAGATGATCAAAACCATGGCCATGAAATATACCA 1183
DB 1397 AAACGTGGTGGTGGAGCCTAATCTCTGGAACAGACACCTGGAGGATTTATGATCA 1456
QY 1184 CTAAGTGTCTACGATACATCAATGGATGATCTCTTTATTAAGAGATGATGGGA 1243
DB 1457 CCATCATGACGACCATCTCTGAATAGTTTTTG-TCAATGGCAGTCGCCCTCAATTGAA 1515
QY 1244 AAGATGGGACTCGTCATGACACGCGCATCTTACTGACTGGTCTGGATATATGTTCTCTG 1303
DB 1516 AGAATGGCAGAGACATGATGATGCAATCTTACTAAGAGATTTGATATTTGTTCTTGA 1575
QY 1304 AGAATGAGCCCTGTGACACTTTGGGATTTGCAACCATAGTGAATGTGTAGTAATATC 1363
DB 1576 AGAATGAACCATGTGACACTCTAGGCTTTGCCCATCATGATGATGCTCTAGTACC 1635
QY 1364 GCAGCTGCACGATTAATGAGATACAGGTCTTGGACTGGCTTCAACATGGCCATGAT 1423
DB 1636 GAAGTTGTACCATCAATGAGGACACAGGACTTGGCCCTTGCCTTCAACATCGCTCATGAT 1695
QY 1424 CTGGACACAACTTTGGCATGATTCATGATGGAAGGAAACATGTGTAAGTCCGAGG 1483
DB 1696 CAGGCGACAACTTTGGTATGATTCACGACGAGAGAGGAAATCCCTGCGAAGAGCTGAAG 1755
QY 1484 GCACATCATGTCCCTCATATGGCAGACGCAATGGAGTCTTCTCTGGTCAOCCCTGCA 1543
DB 1756 GCAATATATGCTCTCCACACTGACCGGAAACATGAGTGTCTTCTGCTCTTCTGCA 1815
QY 1544 GCGGCGAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCCTTCTGATCAG 1603
DB 1816 GCGGCGAGTATCTCAAGAAATTCCTCAGACACACCTCAGCGGGGTGTCTAGTGATGAG 1875
QY 1604 CAAAGCTGTGAAGAAATACAGTATCTGTGAATATCCAGGAGAAATATATATGACAA 1663
DB 1876 CAAAGCAAGCAGGACAGTATTAATATCCGACAACTACAGGACAGATTTATGATGCTG 1935
QY 1664 ACACAGTGGAGTGGAGTTCGGAGAGAGCAAGCAAGCTCTGATGCTGACCTTAAAA 1723
DB 1936 ACACAGTGTAAATGGCAATTTGGAGCAAAAGCCAGTTATGACGCTTGTGTTTGTGA 1995
QY 1724 AGGACATCTGAAGCCCTGTGGTGGCCATCGTATTGGAGAAATATGAGACTAAATTA 1793
DB 1996 AGGATATTTGCAATCATCTTTGGTGGCCACCGAGTAGGCCACAGAGTGTGAGACCAAGTTTA 2055

RESULT 7

US-10-217-774-1

; Sequence 1, Application US/10217774
; Publication No. US20020193583A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020193583A1el Human Proteases and Polynucleotides Encodi
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774

QY 1784 TGCCAGCAGCAGAAAGGACAAATTTGTGGCATGACATGTGTGGCGGAGGACAGTGTG 1843
DB 2056 TGCCCGCAGCAGAAAGGACCGTTTGTGGCTTGTAGTATGTGTGTGCGCAAGCCAGTGG 2115
QY 1844 TGAATATGTGTATGAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 1903
DB 2116 TAAAGTTTGGGAGAGCTCGGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 2175
QY 1904 GTTCCCATGTCTCAGGACCTCGCGAGGGGAGATATCTCATAGGAGTGCCTCTGCAACA 1963
DB 2176 GTTCAAGATGTCTCCGACATGTGTGTGAGAGATCAAGTTCCAGAGAGACACACGCA 2235
QY 1964 ACCCAAGCCATGCGATGGAGGAGTTCTGTGAGGGCTCCACTGCGACTCTGAGAGTCT 2023
DB 2236 ACCCAAGCCATGAGTATGTGTGGCTTATTTCTGTCCAGGTTCTAGCCGATTTATCAG 2295
QY 2024 GCAACAGTCAGAAATGTCCCGGACAGTGTGTGATCTCCGTGCTGCTCAGTGTGCGG 2083
DB 2296 GCAATATTAAACCTTGCATGAAATAGCTTGGATTTTTCGGGCTCAACAGTGTGAGA 2355
QY 2084 ACAACAGCAGACGATTCAGAGGCGGCGACTACAAAGTGAAGCCCTTACACTCAAGTAGA 2143
DB 2356 ATACAGCAACCTTTCGTGGATGTCTTACCAAGTGAAGCCCTTACAAAGTGAAG 2415
QY 2144 ATCAGGACTTATGCAAACTCTACTGTATGCGAGAGGATTTGATTTCTTTCTTTTGT 2203
DB 2416 AGGAAGATCGATGCAAACTGTACTGCAAGGCTGAGAACTTTGAAATTTTTCGAAATG 2475
QY 2204 CAATAAAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATG 2263
DB 2476 CCGCAAGTGAAGATGGAACCTCTCTGCTCCCCAAACAAATATGATTTTGTATTGAG 2535
QY 2264 GGATATGTGAGAGATTTGGATGTGACAAATGCTTGTGATCTGATCTGTGGAAGACGCT 2323
DB 2536 GGGTTTGTGAATAGTGGATGTGATCATGAACTAGGCTCTAAAGCAGTTTCAGATGCTT 2595
QY 2324 GTGGGTGTGTAAAGGAAATTAATCAAGCTGACAGTTCACAGGGGTCTTACACCAAGC 2383
DB 2596 GTGGCTTTGCAAGGTGATAATTCAACTTGAAGTTTATTAAGGCTGTACCTCAAC 2655
QY 2384 ACCACCAACCAACAGTATTATCAGATGTCACATGTCACCTTCTCTGAGCGCGGATAC 2443
DB 2656 AGCAATGAACAAATATATTCGGGTGTCTCATTCAGCTGCGCGCGCGAGAGATCG 2715
QY 2444 GCATCTATGAATGAAGCTCTCTACCTCTCTACATTTCTGTGCGCAATGCGCTCAGA 2503
DB 2716 AATCCAGGAGCTGACAGTTTCTCTCAGTTTACCTCGAGTTTCGAGCTCAGTCAAG 2775
QY 2504 ACTACCTGATGGCAGTGGACCGTGGACTGGCCCGCGCGGTACAAATTTTGGGCACTA 2563
DB 2776 ATTACCTCAAGCGGGCTGGAGCATCGACTGGCTTGGGAGTTTCCCTTGGGAGCA 2835
QY 2564 CTTTCCAGTACAGAGCTCTTATAATGAGCCCGAGAACTTAATCGTACTGGACCAACA 2623
DB 2836 GGTGTAATACAGCGCTCTTTCAACCGCCCGGAACTGTGACCGCCAGGCGCCCA 2895
QY 2624 ACAGACACTGATTTGGA 2642
DB 2896 ATGAGAGCTGTGCTTTGA 2914

RESULT 7

US-10-217-774-1

; Sequence 1, Application US/10217774
; Publication No. US20020193583A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020193583A1el Human Proteases and Polynucleotides Encodi
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774

/ CURRENT FILING DATE: 2002-08-12
 / PRIOR APPLICATION NUMBER: US/09/930,872
 / PRIOR FILING DATE: 2001-08-14
 / PRIOR APPLICATION NUMBER: US 60/225,852
 / PRIOR FILING DATE: 2000-08-16
 / NUMBER OF SEQ ID NOS: 5
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 1
 / LENGTH: 1476
 / TYPE: DNA
 / ORGANISM: homo sapiens
 / US-10-217-774-1

Query Match 22.9%; Score 663; DB 13; Length 1476;
 Best Local Similarity 71.5%; Pred. No. 2.7e-199;
 Matches 1079; Conservative 0; Mismatches 185; Indels 245; Gaps 6;

QY 202 GCAGGACCTGGTGGCCATGGACCCGACGCGCAGCGCTGGGAGCCGAGCGTCCC 261
 DB 69 GCAGGACCTGGTGGCCATGGACCCGACGCGCAGCGCTGGGAGCCGAGCGTCCC 128
 QY 262 GCGTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAAGCGCGGACATGGTGA 321
 DB 129 GCGTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAAGCGCG----- 175
 QY 322 GCTGGAACCATGCTTCTCAGCAAACTAACACAGGAACAGAAACCAAAACACTGCATGT 381
 DB 176 ----- 175
 QY 382 CTCACCTAATATGACCTGTCTCTGCTACGAGTTGACCAAGGGGGGATTAAGTGTCC 441
 DB 176 -----AATATGACCTGTCTCTGCTACGAGTTGACCAAGGGGGGATTAAGTGTCC 228
 QY 442 CATGAATCATGACCATCAGCGCGGAGGAGGAGTGGCGGTGCGAGGTTGAGTCT 501
 DB 229 CATGAATCATGACCATCAGCGCGGAGGAGGAGTGGCGGTGCGAGGTTGAGTCT 280
 QY 502 CTTACACCTTCGGTGAAGGCCGCCAGGACGACTTCCACATGGATCTGAGGACTTCCAGC 561
 DB 289 CTTACACCTTCGGTGAAGGCTCCAGGACGACTTCCAGTGTGATCTGAGGACTTCCAGC 348
 QY 562 AGCTAGTGTCTCTGCTTTATTTGTCAGAGCTTGGGAAGACAGGACCTAAGTCTGTG 621
 DB 349 AGCTAGTGTCTCTGCTTTATTTGTCAGAGCTTGGGAAGACAGGACCTAAGTCTGTG 408
 QY 622 CAGACTTTACCGCAGAGGACTTCTGTTTCTATCAAGGCTCTTGGATFACACAGAAAC 681
 DB 409 CAGACTTTACCGCAGAGGACTTCTGTTTCTATCAAGGCTCTTGGATFACACAGAAAC 468
 QY 682 TCGCCATCGATGAGGAGGAGTCTGTGAGGCGTCCACTCGCACTCTGAGCTCTGCAAC 741
 DB 469 TCCTCAGTGG-----CCCTTCAACCTGCGCAAGGCTGTGCA 504
 QY 742 AGTCAGAAATGTCGCGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCGGAGCACAAC 801
 DB 505 GGATGATGACACAGAGGAGGAGTACTTCTTAAGGCCACTTCTTTCACACCTCTCA 564
 QY 802 AGCAGAGATTCAGAGGCGGACACTACAGTGG--AAGCTTACACTAAGTAGAAGCGG 859
 DB 565 TGGAAATCGGCGAGGTGCCAAGGCGAGCTGCCATCCCACTACTGTACAAAGATCC 624
 QY 860 ACTTATCAAACTACTACTGTATCGCAGAGGATTTGATTTCTTCTTTCTTTCTTCTCAATA 919
 DB 625 ACAGAGCCCATGCTCTCTGGGCGAGTGAAGTCTGTGACCTC-----AAGG 672
 QY 920 AAGTCAAAAGATGGAGCTTCCATGCTCGGAGGATAGCGGTATGTTGTATAGATGGATAT 979
 DB 673 ACATGGGAGCTGGACATCAACCCCTGCACAGAGGAGCCTTCCGCTGGGAGTGCACAA 732
 QY 980 GTGAGCTCAGTGTGGTGTCCATCTCGCG-ACATGCCCGAGCTCCCAAGAGACCTTC 1038
 DB 733 AAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGAGCTCCCAAGGAAGACCTTC 792

QY 1039 TTCATCTTCCGACATGAGTATTAAGTCTTTCACGCAATAAGCGCTCTCTTCTGAGGTCC 1098
 DB 793 TTCATCTTCCGACATGAGTATTAAGTCTTTCACGCAATAAGCGCTCTCTTCTGAGGTCC 852
 QY 1099 CATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTGGTGCACAAAAGATGATGCAA 1158
 DB 853 CATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTGGTGCACAAAAGATGATGCAA 912
 QY 1159 AACCATGGCATGAAATATACACCTACCTACCTGCTCAGCATCTCAACATGATGATCTCT 1218
 DB 913 AACCATGGCATGAAATATACACCTACCTACCTGCTCAGCATCTCAACATGATGATCTCT 972
 QY 1219 TTATTCAAAAGA----- 1229
 DB 973 TTATTCAAAAGATGAACATAGGAGGAACATCAACATTTGCAATTTAGTCTGATCTCT 1032
 QY 1230 ----- 1229
 DB 1033 CTAGAAGATGAACAGCCAGGACTGGTGTATGATGATGATGATGATGATGATGATGATG 1092
 QY 1230 -----TGGATTTGATGGGAAAGATGGGACTCGTCAATGACCAAGCAGCAGC 1272
 DB 1093 TTCTGCGAGTGGGAGTCTGATTTGATGGGAAAGATGGGACTCGTCAATGACCAAGCAGC 1152
 QY 1273 TTACTGACTGCTGGATATATGTTCTCGAAAGATGAGCCCTGTGACACTTTGGGATTT 1332
 DB 1153 TTACTGACTGCTGGATATATGTTCTCGAAAGATGAGCCCTGTGACACTTTGGGATTT 1212
 QY 1333 GCACCCATAGTGGATGTTAGTAAATATCGCAGCTGACGATTAATCAAGATACAGCT 1392
 DB 1213 GCACCCATAGTGGATGTTAGTAAATATCGCAGCTGACGATTAATCAAGATACAGCT 1272
 QY 1393 CTTGGACTGCGCTTCCACCATGCTCCCATGATCTGGAACAACTTTGGCATGATTCATGAT 1452
 DB 1273 CTTGGACTGCGCTTCCACCATGCTCCCATGATCTGGAACAACTTTGGCATGATTCATGAT 1332
 QY 1453 GGAAGAGGAAATGTTGTTAAAGTCCGAGGGGCAACATATGTCCTCCCTACATTTGGCAGGA 1512
 DB 1333 GGAAGAGGAAATGTTGTTAAAGTCCGAGGGGCAACATATGTCCTCCCTACATTTGGCAGGA 1392
 QY 1513 CGCATGAGTCTTCTCTGCTGCTCAGCGCGGAGTATCTACACAAATTTCTTAAGC 1572
 DB 1393 CGCATGAGTCTTCTCTGCTGCTCAGCGCGGAGTATCTACACAAATTTCTTAAGC 1452
 QY 1573 ACGCTTCAA 1581
 DB 1453 TCAGTGAAA 1461

RESULT 8

US-10-029-386-8689/c
 ; Sequence 8689, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: AEOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 8689
 ; LENGTH: 525
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC022424.3
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6


```

; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 65577, A Human Matrix
; FILE REFERENCE: MP101-049P1RNM
; CURRENT APPLICATION NUMBER: US/10/103.377C
; PRIORITY FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/278,347
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2055)
US-10-103-377C-3

Query Match          9.2%; Score 265.8; DB 14; Length 2052;
Best Local Similarity 62.0%; Pred. No. 5.5e-73;
Matches 420; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

QY 1966 CCCAAGCCATCGCATGGAGGAAGTTCTGTGAGGGCTCCACTCGCACTCTCAAGCTCTGC 2025
DB 1321 CACAAGCCCTCAGTATGGTGGCTTATCTGTCCAGGTTCTAGCGTATTTATCAGCTGTC 1380

QY 2026 AACAGTCAGAAATGTCGCCGGACAGTGTTCAGTTCGGTCTGCTCAGTGTGGCGACAC 2085
DB 1381 AATATTAAACCTTGCATGAAATAGCTTGGATTTTCGGGCTCAACAGTGTGAGAATAT 1440

QY 2086 AACACGACAGGATTCAGAGGGCGGCACATACAAAGTGGAGCCCTTACCTCAAGTAGAAGAT 2145
DB 1441 AACACGAAACCTTTCGGTGGATGTTCTACCGTGGAGACCTTACAAAGTGGAGAG 1500

QY 2146 CAGGACTTATGCAAACTCTACTGTATGCAAGAGATTTGATTTCTTCTTTCTGCA 2205
DB 1501 GAAGATCGATGCAAACTCTACTGTGCAAGGCTGAGAACTTTGAAATTTTTTTTGAATGTC 1560

QY 2206 AATAAGTCAAAAGTGGAGCTCCATGCTCGGAGGATAGCCGCTTATGTTGATAGATGG 2265
DB 1561 GCAAGTCAAAAGTGGAGCTCCCTGCTCCCAACAAACAAATGATGTTGATGAGCGG 1620

QY 2266 ATATGTGAGAGTTGGATGTGACAAATGCTCTTGGATCTGATGCTGTGTAAGAGCTGT 2325
DB 1621 GTTGTGAACCTAGTGGGATGTGATCAATGAACCTAGGCTTAAAGCAGTTTCAGATGCTGT 1680

QY 2326 GGGGTGTCTAAGGGAATAACTCAGCCTGCACGATTCACAGGGTCTCTACCAAGCAC 2385
DB 1681 GCGTTTGCAAAGTGTGATATCAACTTGCAGTTTATTAAGGCTTGTACCTCAACAG 1740

QY 2386 CACACACCAACCAAGTATATACATGCTACCAATTCCTTCTGGAGCCCGAGTATCCG 2445
DB 1741 CATAAAGCAAAATGATATATATCCGGTGGTCCCTTATTCAGCTGGCGCCGAGCATCGAA 1800

QY 2446 ATCTATGAATGAAGCTCTCTACCTCTACATTTCTGTGGCGCAATGCCCTCAGAAAGTAC 2505
DB 1801 ATCCAGGAGCTGCAGGTTTCTCCNGTTTACCTCCAGTTTCAGACCTCAGTCAAAAGTAT 1860

QY 2506 TACCTGAATGGGCACTGGACCGTGTGGCCCGCGGTACAAATTTTCGGGCACTACT 2565
DB 1861 TACCTCAGCGGGGCTGGAGCATCGACTGGGCTGGGAGTTCCCTTTCGCTGGGACACG 1920

QY 2566 TTCGACTACAGAGCTCTCTATTAAGTGGCCCGGAGACTTAATCGTACTGAGCAACCAAC 2625
DB 1921 TTGTAATACCAAGCTCTTTCAACCGCCCGGAAACGCTCTGTACCGCCGAGGCGCCACAAAT 1980

QY 2626 GAGACACTGATTGTGGA 2642
DB 1981 GAGACGCTGGTCTTTGA 1997

RESULT 12
US-10-103-377C-1
; Sequence 1, Application US/10103377C
; Publication No. US20030073098A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 65577, A Human Matrix
; FILE REFERENCE: MP101-049P1RNM
; CURRENT APPLICATION NUMBER: US/10/103.377C
; PRIORITY FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/278,347
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (323)...(2374)
US-10-103-377C-1

Query Match          9.2%; Score 265.8; DB 14; Length 2377;
Best Local Similarity 62.0%; Pred. No. 6.1e-73;
Matches 420; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

QY 1966 CCCAAGCCATCGCATGGAGGAAGTTCTGTGAGGGCTCCACTCGCACTCTCAAGCTCTGC 2025
DB 1643 CACAAGCCCTCAGTATGGTGGCTTATCTGTCCAGGTTCTAGCGTATTTATCAGCTGTC 1702

QY 2026 AACAGTCAGAAATGTCGCCGGACAGTGTTCAGTTCGGTCTGCTCAGTGTGGCGACAC 2085
DB 1703 AATATTAAACCTTGCATGAAATAGCTTGGATTTTCGGGCTCAACAGTGTGAGAATAT 1762

QY 2086 AACACGACAGGATTCAGAGGGCGGCACATACAAAGTGGAGCCCTTACCTCAAGTAGAAGAT 2145
DB 1763 AACAGCAAAACCTTTCGGTGGATGTTCTACCAAGTGGAAACCTTATACAAAGTGGAGAG 1822

QY 2146 CAGGACTTATGCAAACTCTACTGTATGCAAGAGATTTGATTTCTTCTTTCTGCA 2205
DB 1823 GAAGATCGATGCAAACTCTACTGTGCAAGCTGAGAACTTTGAAATTTTTTTTGAATGTC 1882

QY 2206 AATAAGTCAAAAGTGGAGCTCCATGCTCGGAGGATAGCCGCTTATGTTGATAGATGG 2265
DB 1883 GCAAAAGTCAAAAGTGGAGCTCCCTGCTCCCAACAAACAAATGATGTTGATGAGCGG 1942

QY 2266 ATATGTGAGAGTTGGATGTGACAAATGCTCTTGGATCTGATGCTGTGTAAGAGCTGT 2325
DB 1943 GTTGTGAACCTAGTGGGATGTGATCAATGAACCTAGGCTTAAAGCAGTTTCAGATGCTGT 2002

QY 2326 GGGGTGTCTAAGGGAATAACTCAGCCTGCACGATTCACAGGGTCTCTACCAAGCAC 2385
DB 2003 GCGTTTGCAAAGTGTGATATTAATTCAGCTTGAAGTTTATTAAGGCTTGTACCTCAACAG 2062

QY 2386 CACACACCAACCAAGTATATACATGCTACCAATTCCTTCTGGAGCCCGAGTATCCG 2445
DB 2063 CATAAAGCAAAATGATATATATCCGGTGGTCCCTTATTCAGCTGGCGCCGAGCATCGAA 2122

QY 2446 ATCTATGAATGAAGCTCTCTACCTCTACATTTCTGTGGCGCAATGCCCTCAGAAAGTAC 2505
DB 2123 ATCCAGGAGCTGCAGGTTTCTCCAGTTTACCTCCAGTTTCAGACCTCAGTCAAAAGTAT 2182

QY 2506 TACCTGAATGGGCACTGGACCGTGTGGCCCGCGGTACAAATTTTCGGGCACTACT 2565
DB 2183 TACCTCAGCGGGGCTGGAGCATCGACTGGGCTGGGAGTTCCCTTTCGCTGGGACACG 2242

QY 2566 TTCGACTACAGAGCTCTCTATTAAGTGGCCCGGAGACTTAATCGTACTGAGCAACCAAC 2625
DB 2243 TTGTAATACCAAGCTCTTTCAACCGCCCGGAAACGCTCTGTACCGCCGAGGCGCCACAAAT 2302

QY 2626 GAGACACTGATTGTGGA 2642
```


1118 TTCCAGCAGCTGAGGAGACACTGTGTCAAACTGGATATTTGAAAGGGTGTGTATC 1177
1844 TGAATATGTTGAT-----GAAGCCCAAGCCCACTGCGGCTGTGTGACT 1894
1178 AGGAGATGTTGTTCTTTGGCACTTGGCCCAAGACATAGATGGGGCTGGGCTCCT 1237
1895 GGTCTTCTGTCCTCCATGTCAGGACCTGCGGAGGGAGATCTCATAGAGTGGCC 1954
1238 GGTCACTATGGGAGAGTGCAGCAGGACCTGCGGGGAGGCGTCTCTCATCCCTAAGAC 1297
1955 TCTGACCAACCCCAAGCCATCGATGGAGGAGTCTCTGAGGGCTCCACTCGCACTC 2014
1298 ACTGTGACGTTCAGCAGCTTCAGAGGTGGAATAATATTCCTTGGGAGAGAAAGGT 1357
2015 TGAAGCTCTGCAACAGTCAGAAATGTCCTCCGAGAGTGTGACTTCCGTGCTCAGT 2074
1358 ATCGCTCTGTAAACAGATCCATGCTTGGGTTCCCGAGATTTTCGAGAGAAACAGT 1417
2075 GTGCGGACCAACAGCAGAGCTTCAGAGGCGGCACTACAGTGGAGCCTTACACTC 2134
1418 GTGAGACTTTGACAAATATGCTTTCGAGAAAGTATTAATCTGGAACCTTATACT- 1476
2135 AAGTAGAAGTACAGACTTATGCAAACTCTACTGTATCGCAAGAGATTTGATTTCTCT 2194
1477 --GGAGGTGGGTAAACCTTGTGCAATTAACCTGCTGGCTGAAGGTATTAATTTCTACA 1534
2195 TTTCTTTGTCAATAAAGTCAAGATGGGACTCATGCTCGGAGGATAGCCGTAATGTTT 2254
1535 CTGAAGCTGCTCTGCGGTGATCGATGGGAGCCCAAGTGCATGCGGATCTCATGATATCT 1594
2255 GTATAGATGGGATATGTGAGAGAGTGTGATGTGCAATGTCCTTGGATCTGATGCTGTG 2314
1595 GCATCAATGGAGATGCAAGCAGCTAGGCTGTGATATATTTTGGGATCTGATGCTAGGG 1654
2315 AAGAGCTGTGGGTGTGTAACCGGAATAACTCAGCTGCAAGTTCACAGGGTCTCT 2374
1655 AAGATAGATGTGAGTCTGTGGAGGAGCGGAACATGTGATGCCATTTGAAGGGTCT 1714
2375 ACACCAAGCAGCAGCACCAACCAAGTATATCATATGTCACATTCCTTCTTGGAGCCC 2434
1715 TCAATGATTCCTGCCAGGGAGGCTACATGGAAGTGTGCGATACCAAGAGGCTCTG 1774
2435 GGAGTATCCGATCTATGAATGAACGCTCTACTCTCTACATTTCTGCGGCAATGCC 2494
1775 TTCATTTGAAGTTAGAGAGTTCCTCATGTCAAGAACTATATGCTTTAAATCTGAAG 1834
2495 TCAGAGGCTACTCTGAAATGGGCACTGGACCGTGGACTGGCCGCGCGGTACAAATTT 2554
1835 GAGATGATTAATATTAATGCTGCTGGACTATTTGCTGGCTAGGAATTTGATGTTG 1894
2555 CGGGCACTATTTCGACTAGACGGTCTTATATAGCCGCGAGAACTTAATCGCTACTG 2614
1895 CTGGACAGCTTTTCAATTACAGAGACCAACTGATGAACAGAACTCTTGAAGAGCTTAG 1954
2615 GACCAACCAAGCAGACACTGATTTGGAGTGTCTTTTTCAGGAGGAAGAACCGGGTGT 2673
1955 GTCCTACCTCAGAAATCTCATGCTGATGCTGCTTCTGCTTCAGAAACAGAAATTTGGAAAT 2013

RESULT 15

US-09-963-791-1

; Sequence 1, Application US/09963791

; Patent No. US20020120113A1

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Scoville, John

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. US20020120113A1 Human Proteases and Polynucleotides Encodir

; FILE REFERENCE: LEX-0105-USA

; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2727
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-791-1

Query Match 7.8%; Score 227; DB 10; Length 2727;
Best Local Similarity 50.1%; Pred. No. 1.4e-60;
Matches 721; Conservative 0; Mismatches 685; Indels 33; Gaps 5;
QY 1253 CTCGTCTATGACCAAGCCCATCTTACTGCTGTGTGATATATGTTCTCTGAAAGATGAGC 1312
DB 1043 CCCACCAGATAATGCAGTTCTTATTACTAGATATGATATCTGCACCTTATAAAATAAGC 1102
QY 1313 CTTGTGACACTTTGGGATTTGCACCATAAGTGAATGTGTAGTAATATCGCAGCTGCA 1372
DB 1103 CTTGTGGAACACTGGGCTTGGCCCTCTGTGGCTGGAATGTGTGAGCCCTGAAAGGAGCTGCA 1162
QY 1373 CGATTAAATGAAGATACAGCTCTTGGACTTGGCCCTTACCATTTGCCATTCAGCTGAGTGGACACA 1432
DB 1163 GCATTAATGAAGACATTTGGCTTGGCTTTCAGCTTTTACCATTTGCACATGAGATTTGTCACA 1222
QY 1433 ACTTTGGCATGATTCATGATGGAGAGGAAACATGTGTATAAA-----AGTCCGAGG 1483
DB 1223 ATTTTGGTATGAACCATGATGGAATTTGGAAATTTCTTGGGAGCAAGAAAGGCTCATGAAGCAG 1282
QY 1484 GCACATCATCTGCCCTACATTTGGCAGAGCGCATGGAGTCTTCTCTGCTGACCCCTGCA 1543
DB 1293 CAAACTTATGGAGCTCACATTTACTGGATACCAATCTTTTCTGCTGCTGCTGCA 1342
QY 1544 GCGGCACTATCTACAAATTTCTAAGACCGCTCAAGCTATCTGCTTCTGCTGATCAGC 1603
DB 1343 GTGAGACTACATCACCAGCTTTCTAGATTTCAGGCGGTGGTACTTGGCTTGTATAATGAGC 1402
QY 1604 CAAAGCTGTCAAGCAATACAAATCTCTGAGAAATTTGCCAGGAGAAATTTATATGATCAAA 1663
DB 1403 CT---CCCAAGGCTGACTTTCTTTATCCAGCTGTGGCCCAAGCTCAGTGTATGATGCTG 1459
QY 1664 ACACACAGTGCAGTGGCAGTTTCGGAGAGAAAGCAAGCTCTGCATGCTGGACTTTTAAA 1723
DB 1460 ATGACCAATGTCTTTCCAGTATGGAGCAACTCCCGCAATGTAAATATG----- 1510
QY 1724 AGGACATCTGTAAAGCCCTGTGTGCCATCTTATTTGGAGGAATGTGAGACTTAATTTA 1783
DB 1511 GGGAAAGTGTGTAGAGAGCTCTGGGTGTCTCAGCAAAAGCAACCGCTGTGTGCCAAACAGTA 1570
QY 1784 TGCCAGCAGCAGAGAGCAATTTTGTGGCATGATGATGTGTGCGGGGAGGAGCAGTGTG 1843
DB 1571 TTCCAGCAGCTGAGGGAGACTGTGTCAAACTGGGAATATTGAAAAGGGTGGTGTATC 1630
QY 1844 TGAATATGCTGAT-----GAAGCCCCCAAGCCCACTGAGCCACTGGTGGACT 1894
DB 1631 AGGGAGATGTGTCTTTTGGCACTTTGGCCCTCAGAGCATAGATGGGGGCTGGGGTCCCT 1690
QY 1895 GGTCTTCTTGGTCCCAAGCTTCCAGGACCTGCGGAGGGGAGTATCTCATAGAGTGGCC 1954
DB 1691 GGTCACTATGGGAGAGTGCAGCAGAGACTTGGGGGAGGCGCTCTCTCATCTCCCTAAGAC 1750
QY 1955 TCTGCAACCAACCCCAAGCCATCGCATGAGAGGAGATTTCTGTGAGGGCTCCACTCGCAGCTC 2014
DB 1751 ACTGTGACAGTCCAGACACTTTCAGGAGGTGGAAAAATATTGCTTTGGGAGAAAGAAAGGT 1810
QY 2015 TGAAGCTCTGCAACAGTCAAGAAATGTCCCGGAGACAGTGTGTGACTTCCGCTGCTCAGT 2074
DB 1811 ATCGCTCTGTAAACAGATCCATGCTTGGGTTCCCGAGATTTTCCGAGAAACAGT 1870
QY 2075 GTGCGGAGCACACAGCAGCAGATTCAGAGGGGGCACTACAAAGTGGAAAGCCCTTACACTC 2134

Wed Oct 29 15:41:41 2003

```

||||| 1871 GTCCAGACTTTGCAATATGCTCTTCGAGGAAAGTATTATTAACCTGGAACCCCTATACT- 1929
Qy
2135 AAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAAAGGATTGATTTCTTCT 2194
Db
1930 --GGAGGTGGGTAAACCTTGTGCATTAAACTGCTTGGCTGAAGGTATATTTCTACA 1987
Qy
2195 TTTCTTTGTCAATAAAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTT 2254
Db
1988 CTGAACGTCTCTCGGGTGTATCGATGGGACCCAGTGCATGCGGATTCACTGGATATCT 2047
Qy
2255 GTATAGATGGATATCTGAGAGAGTTGGATGTGCAATGTCCTTGGATCTGTATGCTGTTG 2314
Db
2048 GCATCAATGAGAAATGCAACGCTAGGCTGTGATATATTTGGGATCTGATGCTAGGG 2107
Qy
2315 AAGACGTCTGTGGGGTGTAAACGGGAATAACTCAGCCTGCACGATTTCACAGGGGTCTCT 2374
Db
2108 AAGATAGATGTCGAGTCTGTGGAGGGGACGGAAGCACATGTGATGCCATTGGAAGGTCT 2167
Qy
2375 ACACCAAGCACCCACACCAACCGATATTATCACATGGTCACCATTCCTTCTGGAGCCC 2434
Db
2168 TCAATGATTCACCTGCCCGGAGGGCTACATGGAAGTGGTGCAGATACCAAGAGGCTCTG 2227
Qy
2435 GGAGTATCCGCATCTATGAATGAACGCTCTCTACTCTCTACATTTCTGTGCGCAATGCC 2494
Db
2228 TTCACATGAAAGTTAGAGAGTTGCCATGTCAAGAACTATATGCTTTAAATCTGAAG 2287
Qy
2495 TCAGAAAGTACTACTCTGAATGGGCACTGGACCGTGGACTGGCCCGGCGGTACAAATTTT 2554
Db
2288 GAGATGATTACTATATTAAATGGTGGCTGGACTATTGACTGGCCCTAGGAAATTTGATGTTG 2347
Qy
2555 GGGCACTACTTTCGACTACAGCGGTCTTATATGAGCCCGGAGACTTAAATCGCTACTG 2614
Db
2348 CTGGGACAGCTTTTCATTACAGAACCAACTGATGAACCAAGATCCCTTGGAAAGCTCTAG 2407
Qy
2615 GACCAACCAACGAGACACTGATTTGGAGTGTCTGTTTCAGGGAAGGAACCCGGGTGTT 2673
Db
2408 GTCCTACTCAGAAATCTCATCGTCATGCTTCATGCTTCAAGAACAGANTTGGGAATT 2466
```

Search completed: October 28, 2003, 23:39:24
Job time : 764 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 19:02:54 ; Search time 82 Seconds
(without alignments)
1842.777 Million cell updates/sec

Title: US-09-981-151a-8

Perfect score: 5236

Sequence: 1 MKPRKRWGLAALMLLAQ.....LEACQPSATVIALAFLES 952

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107663 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107663

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

```
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	5236	100.0	952	23	ABG76897 Human ADAM-TS 7-li
2	5110	97.6	986	23	ABG76894 Human zinc metallo
3	4103.5	78.4	791	23	ABG76895 Human zinc metallo
4	3858.5	73.7	1224	23	AAU79497 Human ADAMTS prote
5	3858.5	73.7	1224	23	AAU72891 Human metalloprote
6	3846.5	73.5	1224	23	AAO15254 Human MDT9 protea
7	3846.5	73.5	1224	24	ABG72431 Novel human protea
8	3840.5	73.3	1224	24	AAU80153 Human ADAMTS prote
9	3750.5	71.6	1170	23	AAU79500 Human ADAMTS prote

10	3365	64.3	856	23	ABG76896	Human ADAM-TS 7-li
11	3286	62.8	1236	23	AAU98886	Human protease PR
12	3280	62.6	1021	23	AAU79496	Human partial ADAM
13	2812.5	53.7	556	23	ABG71503	Human ADAM-TS-like
14	2435.5	46.5	1221	23	AB881460	Human aggrecanase
15	2364	45.1	1162	24	ABJ19379	NOVX related prote
16	2265	43.3	862	24	ABR38957	Human protease (NH
17	2196.5	41.9	934	23	ABU72893	Human metalloprote
18	1543.5	29.5	491	24	ABG72430	Novel human protea
19	1391	26.6	959	22	AA86947	Human metalloprote
20	1390	26.5	908	22	AAE03572	Human protease-rel
21	1380.5	26.4	1120	23	ABG76502	DNA encoding prote
22	1342	25.6	353	24	ABJ19382	NOVX related prote
23	1342	25.6	353	24	ABJ19384	NOVX related prote
24	1339	25.6	353	24	ABJ19381	NOVX related prote
25	1332	25.4	349	24	ABJ19380	NOVX related prote
26	1332	25.4	353	24	ABJ19383	NOVX related prote
27	1319	25.2	757	22	AAE03583	Human protease-rel
28	1257	24.0	1690	22	AA86949	Human metalloprote
29	1247	23.8	1686	22	AAE00913	Human 27875 ADAM-T
30	1246	23.8	1686	22	AAE00913	Human ADAM type me
31	1246	23.8	1686	22	AAE00934	Amino acid sequenc
32	1243	23.7	1044	23	AB878141	Human metalloprote
33	1243	23.7	1044	23	AB807574	Human aggrecanase
34	1243	23.7	1104	23	AAU97888	DNA encoding prote
35	1242	23.7	1103	23	ABG76505	Human metalloprote
36	1242	23.7	1103	23	AAU72890	Human ADAM type me
37	1240	23.7	1103	22	AA874945	Human ADAM type me
38	1240	23.7	1104	23	AA847719	ADAMTS-E. Homo sa
39	1234	23.6	1133	23	AAE19375	Human 33428 protei
40	1231.5	23.5	1132	23	AB882027	Human metalloprote
41	1229	23.5	1103	23	ABP69312	Human polypeptide
42	1222.5	23.3	1068	23	AB878136	Amino acid sequenc
43	1222.5	23.3	1068	23	AB807572	Human metalloprote
44	1208	23.1	1784	21	AA841379	Human ORFX ORF113
45	1207	23.1	852	24	ABU12116	Human protein mod

ALIGNMENTS

RESULT 1

ABG76897
ID ABG76897 standard; Protein; 952 AA.
XX
AC ABG76897;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human ADAM-TS 7-like protein #2.
XX

Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing;
breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
autoimmune disease; allergy; addiction; asthma; transplantation;
graft versus host disease; systemic lupus erythematosus; scleroderma;
psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
thrombocytopenia; bleeding disorder; metabolic disorder; obesity;
glucose transport defect; glomerulonephritis; hypercalcaemia;
polycystic kidney disease; renal tubular acidosis; skin disorder;
congenital diarrhoea; respiratory disease; gastro-intestinal disease;
muscle disorder; bone disorder; joint disorder; skeletal disorder;
haematopoietic disorder; urinary system disorder; osteoporosis;
dental disease; dental infection; growth disorder; reproductive disorder;
hypogonadism; fertility disorder; viral infection; bacterial infection;
parasitic infection; metabolic pathway modulation; gene therapy;
zinc metalloprotease; ADAM-TS 7; alpha-2-macroglobulin precursor;
ileal sodium/bile acid cotransporter; prohibitin; MT; CIP4; spinesin;
macrophage stimulating protein precursor; fatty acid-binding protein;
gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.

Human zinc metalloprotease-like protein #1.

Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing; breast cancer; Alzheimer's disease; epilepsy; Huntington's disease; anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis; neurodegeneration; Parkinson's disease; pain; stroke; endometriosis; autoimmune disease; allergy; addiction; asthma; transplantation; graft versus host disease; systemic lupus erythematosus; scleroderma; psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus; atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis; thrombocytopenia; bleeding disorder; metabolic disorder; obesity; glucose transport defect; glomerulonephritis; hypercalcaemia; polycystic kidney disease; renal tubular acidosis; skin disorder; congenital diarrhoea; respiratory disease; gastro-intestinal disease; muscle disorder; bone disorder; joint disorder; skeletal disorder; haematopoietic disorder; urinary system disorder; osteoporosis; dental disease; dental infection; growth disorder; reproductive disorder; hypogonadism; fertility disorder; viral infection; bacterial infection; parasitic infection; metabolic pathway modulation; gene therapy; zinc metalloprotease; ADAM-TS 7; alpha-2-macroglobulin precursor; ileal sodium/bile acid cotransporter; prohibitin; Wt; CIP4; spinosin; macrophage stimulating protein precursor; fatty acid-binding protein; gap junction beta-5 protein; hepsin/plasma transmembrane serine protease; single nucleotide polymorphism; SNP.

Human sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

PH Misc-difference 155 /note= "Arg substituted by Gly as a result of a single nucleotide polymorphism (SNP)"

FT FT

FT Misc-difference 158 /note= "Ala substituted by Pro as a result of a single nucleotide polymorphism (SNP)"

FT FT

FT Misc-difference 373 /note= "Val substituted by Ala as a result of a single nucleotide polymorphism (SNP)"

FT FT

FT Misc-difference 553 /note= "Cys substituted by Arg as a result of a single nucleotide polymorphism (SNP)"

FT FT

XX XX

XX W0200233087-A2.

XX XX

XX 25-APR-2002.

XX XX

XX 17-OCT-2001; 2001WO-US32496.

XX XX

XX 17-OCT-2000; 2000US-241040P.

XX XX

XX 17-OCT-2000; 2000US-241058P.

XX XX

XX 17-OCT-2000; 2000US-241063P.

XX XX

XX 17-OCT-2000; 2000US-241243P.

XX XX

XX 20-OCT-2000; 2000US-242152P.

XX XX

XX 23-OCT-2000; 2000US-242482P.

XX XX

XX 23-OCT-2000; 2000US-242611P.

XX XX

XX 23-OCT-2000; 2000US-242612P.

XX XX

XX 24-OCT-2000; 2000US-242888P.

XX XX

XX 24-OCT-2000; 2000US-242881P.

XX XX

XX 29-DEC-2000; 2000US-259028P.

XX XX

XX 20-FEB-2001; 2001US-269813P.

XX XX

XX 25-MAY-2001; 2001US-286324P.

XX XX

XX 29-APR-2001; 2001US-294108P.

XX XX

XX 09-JUL-2001; 2001US-303698P.

XX XX

XX 16-OCT-2001; 2001US-0981151.

XX XX

XX (CURA-) CURAGEN CORP.

XX XX

XX Edinger S, Gerlach JA, MacDougall JK, Malyankar UM, Smithson G, Millet I, Peyman VA, Stone DJ, Gunther E, Ellerman K, Shinkets RA, Padigaru M, Guo X, Patturajan M, Taupier RJ, Burgess CB, Zertusen BD, Kekuda R, Spytek KA, Gangolli EA, Fernandes ER, Gorman L;

XX XX

DR WPI; 2002-590434/63.

DR N-PSDB; ABS59323.

XX Cytoplasmic, nuclear, membrane bound and secreted polypeptides and nucleic acids encoding the polypeptides for diagnosing and treating e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and diabetes.

XX Claim 1; Page 14; 305pp; English.

XX The present invention relates to new NOVX (NOV1-10) polypeptides. The molecules of the invention are useful for treating or preventing a NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or a disorder related to cell signal processing and metabolic pathway modulation in humans. NOVX polypeptides, nucleic acids and antibodies are useful for treating or preventing disorders or syndromes including breast cancer, Alzheimer's disease, epilepsy, Huntington's disease, anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis, neurodegeneration, Parkinson's disease, pain, stroke, autoimmune disease, allergy, addiction, asthma, endometriosis, graft versus host disease, systemic lupus erythematosus, scleroderma, transplantation, psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection, atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes, thrombocytopenia, bleeding disorders, metabolic disorders, obesity, kidney disease, pancreatitis, renal tubular acidosis, skin disorders, congenital diarrhoea, respiratory disease, gastro-intestinal diseases, muscle, bone, joint and skeletal disorders, haematopoietic disorders, urinary system disorders, osteoporosis, dental disease and infection, growth and reproductive disorders, hypogonadism, fertility, and/or other pathologies and disorders, viral, bacterial, or parasitic infections. The present amino acid sequence represents a NOVX protein of the invention.

XX Sequence 986 AA;

XX SQ

Query Match 97.6%; Score 5110; DB 23; Length 986;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 940; Conservative 1; Mismatches 3; Indels 50; Gaps 4;

QY 1 MKPRARGWGLAALWMLLAQVAEQVSPGSHQNRGSGOLEAPPLISGRGPRLTAMS 60
DB 1 MKPRARGWGLAALWMLLAQVAEQVSPGSHQNRGSGOLEAPPLISGRGPRLTAMS 60
QY 61 PLFSAGTCVRHGTRSGSAWEPEPPASSSTGCAAGLDGKGRDMDAGNHRSSQTTGTENQ 120
DB 61 PLFSAGTCVRHGTRSGSAWEPEPPASSSTGCAAGLDGKGRDMDAGNHRSSQTTGTENQ 120
QY 121 TLHVLTVQDLYSAVEVDHGRDYNSHEIMHQRRRAVAVSEVSLHLKXGPRHDFHMDL 180
DB 121 TLHVLTVQDLYSAVEVDHGRDYNSHEIMHQRRRAVAVSEVSLHLKXGPRHDFHMDL 180
QY 181 RTSSSLVAPGFIQVTLGKTGTSVQTLPPDFCFYQGSLSHRNSPSHGKFCGSGTRTL 240
DB 181 RTSSSLVAPGFIQVTLGKTGTSVQTLPPDFCFYQGSLSHRNSPSHGKFCGSGTRTL 240
QY 241 KLCNSQKCPDSVDFRAAQCAEHNSRRPRGRHYKWKPYTQVEADLCKLYCIAGDFDFFS 300
DB 241 KLCNSQKCPDSVDFRAAQCAEHNSRRPRGRHYKWKPYTQVEADLCKLYCIAGDFDFFS 300
QY 301 LSNKVQDTPCSEDSRNVCIDGICELSVSTSAHWPQPKEDLFLPDEYKSLRHKRSL 360
DB 301 LSNKVQDTPCSEDSRNVCIDGICELSVSTSAHWPQPKEDLFLPDEYKSLRHKRSL 360
QY 361 LRSHRNEELNVETLVVDDKKMMQNHGHEINTTVTLTLNMMVSALFKD----- 407
DB 361 LRSHRNEELNVETLVVDDKKMMQNHGHEINTTVTLTLNMMVSALFKDGTGGGINIAVG 420
QY 408 -----GLMGKQDGRHDBHAILLTGLDTCWKNBPCDT 438
DB 421 LILLEDEQGLVISHHADHTLSSFCQWQSGMLGKDGTRHDBHAILLTGLDTCWKNBPCDT 480
QY 439 LGFAPISGMSKYSRCSCTINEDTGLGLAFTTAHSGHNFMIHDEGNGKCKSGNTMSPT 498

Db 481 LGFAPISGCMKSYRSTINEDTGLGLAFTIAHESGHNFMIHDEGNCKSEGNMSPT 540
 QY 499 LAGNGVFSWSPCSOYLHFKFSTAGAICLAQPKVPVKEYPEKLPGLYDANTQCKWQ 558
 Db 541 LAGNGVFSWSPCSOYLHFKFSTAGAICLAQPKVPVKEYPEKLPGLYDANTQCKWQ 600
 QY 559 FGEKAKLMLDFKDKICKALWCHRIGRKCEKFMFAAEGTICGHDMWCRGGQCVKYGDEG 618
 Db 601 FGEKAKLMLDFKDKICKALWCHRIGRKCEKFMFAAEGTICGHDMWCRGGQCVKYGDEG 660
 QY 619 SKTHGHSWSSWSPCSRTCCGGVSHRSLCTNPKPSHGKFCGSTRTLKLSQKCP 678
 Db 661 PKTHGHSWSSWSPCSRTCCGGVSHRSLCTNPNPSHGKFCGSTRTLKLSQKCP 720
 QY 679 RDSVDFRAACAEHNSRRFRGRHYKWKPTQVEDQDLCKLYCIAEGDFDFFLSKNKVDG 738
 Db 721 RDSVDFRAACAEHNSRRFRGRHYKWKP-----QDLCKLYCIAEGDFDFFLSKNKVDG 774
 QY 739 TPCSEDSRNVICDIGICBVCDCNVLGSDAVEDVCGVCNGNSACTIHRGLYTKHHTNQY 798
 Db 775 TPCSEDSRNVICDIGICE-XGCDNVLGSDAVEDVCGVCNGNSACTIHRGLYTKHHTN-H 832
 QY 799 YHVVTPSGARSIRIYEMNVSTSYISVRNALRYLNGHWTVDMPGRYKFSCTTFDYRRS 858
 Db 833 YHVVTPSGARSIRIYEMNVSTSYISVRNALRYLNGHWTVDMPGRYKFSCTTFDYRRS 892
 QY 859 YNPEPNIATGPNETLIVELLFQGRNPGVAMEYSNPRLTGTEKQPPAQPSYTWAIVRSEC 918
 Db 893 YNPEPNIATGPNETLIVELLFQGRNPGVAMEYSNPRLTGTEKQPPAQPSYTWAIVRSEC 952
 QY 919 SVSCGGRCPLVLLLEAACOPSAVYALAFLES 952
 Db 953 SVSCGGRCPLVLLLEAACOPSAVYALAFLES 986

RESULT 3

ABG76895
 ID ABG76895 standard; Protein; 791 AA.
 AC ABG76895;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human zinc metalloprotease-like protein #2.
 XX
 KW Human; NOXV; cardiomyopathy; atherosclerosis; cell signal processing;
 KW breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
 KW anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
 KW neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
 KW autoimmune disease; allergy; addiction; asthma; transplantation;
 KW graft versus host disease; systemic lupus erythematosus; scleroderma;
 KW psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
 KW atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
 KW thrombocytopenia; bleeding disorder; metabolic disorder; obesity;
 KW glucose transport defect; glomerulonephritis; hypercalcaemia;
 KW polycystic kidney disease; renal tubular acidosis; skin disorder;
 KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
 KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
 KW haematopoietic disorder; urinary system disorder; osteoporosis;
 KW dental disease; dental infection; growth disorder; reproductive disorder;
 KW hypogonadism; fertility disorder; viral infection; bacterial infection;
 KW parasitic infection; metabolic pathway modulation; gene therapy;
 KW zinc metalloprotease; ADAM-TS 7; alpha-2-macroglobulin precursor;
 KW ileal sodium/bile acid cotransporter; prohibitin; MT; CIP4; spinosin;
 KW macrophage stimulating protein precursor; fatty acid-binding protein;
 KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.

OS Homo sapiens.

XX WO2000233087-A2.

XX

PD 25-APR-2002.

XX 17-OCT-2001; 2001WO-US32496.
 XX
 PR 17-OCT-2000; 2000US-241040P.
 PR 17-OCT-2000; 2000US-241058P.
 PR 17-OCT-2000; 2000US-241063P.
 PR 17-OCT-2000; 2000US-241243P.
 PR 23-OCT-2000; 2000US-242152P.
 PR 23-OCT-2000; 2000US-242482P.
 PR 23-OCT-2000; 2000US-242611P.
 PR 23-OCT-2000; 2000US-242612P.
 PR 24-OCT-2000; 2000US-242880P.
 PR 24-OCT-2000; 2000US-242881P.
 PR 29-DEC-2000; 2000US-259028P.
 PR 20-FEB-2001; 2001US-269813P.
 PR 25-APR-2001; 2001US-286324P.
 PR 29-MAY-2001; 2001US-294108P.
 PR 09-JUL-2001; 2001US-303698P.
 PR 16-OCT-2001; 2001US-0981151.

(CURA-) CURAGEN CORP.

Edinger S, Gerlach V, MacDougall JR, Malyankar UM, Smithson G;
 Millet I, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shimkets RA;
 Padigar M, Guo X, Patturajan M, Taupier RJ, Burgess CE;
 Zehusen BD, Kekuda R, Spytek KA, Gangolli EA, Fernandes ER;
 Gorman L;

WPI; 2002-590434/63.

DR N-PSDB; ABS59324.

Cytoplasmic, nuclear, membrane bound and secreted polypeptides and
 nucleic acids encoding the polypeptides for diagnosing and treating
 e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and
 diabetes

Claim 1; Page 15; 305pp; English.

The present invention relates to new NOXV (NOV1-10) polypeptides. The
 molecules of the invention are useful for treating or preventing a
 NOXV-associated disorder, such as cardiomyopathy, atherosclerosis, or
 a disorder related to cell signal processing and metabolic pathway
 modulation in humans. NOXV polypeptides, nucleic acids and antibodies
 are useful for treating or preventing disorders or syndromes including
 breast cancer, Alzheimer's disease, epilepsy, Huntington's disease,
 anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis,
 neurodegeneration, Parkinson's disease, pain, stroke, autoimmune
 disease, allergies, addiction, asthma, endometriosis, graft versus host
 disease, systemic lupus erythematosus, scleroderma, transplantation,
 psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,
 atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes,
 thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic
 kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 urinary system disorders, osteoporosis, dental disease and infection,
 growth and reproductive disorders, hypogonadism, fertility, and/or other
 pathologies and disorders, viral, bacterial, or parasitic infections.
 The present amino acid sequence represents a NOXV protein of the
 invention.

SQ Sequence 791 AA;

Query Match 78.4%; Score 4103.5; DB 23; Length 791;

Best Local Similarity 81.2%; Pred. No. 0;

Matches 774; Conservative 9; Mismatches 7; Indels 163; Gaps 8;

QY 1 MKPRARGWGRGIALWMLLAQVAEONSPGRSHQGRNGSGOLEASPRLLSRGPRLTAMS 60

DB 1 MKPRARGWGRGIALWMLLAQVAEONSPGRSHQGRNGSGOLEASPRLLSRGPRLTAMS 60
 1 MKPRARGWGRGIALWMLLAQVAEONSPGRSHQGRNGSGOLEASPRLLSRGPRLTAMS 60

QY 61 PLFSAGTCVRHGRTRSGSANEPRPASSTRGAAGLDKGRDMDEAGNHRSSQQTNTGTENQ 120

Db 61 PLFSAGTCVHRGTRSGSAWEPERPASSSTRGAGLDGKGRDMDGAGNHSQQTNGTENQ 120
 Qy 121 TLHVLVT-QVDLSAYEVDHGRGDYVSHIEMHQRERRRAVAVSEVESHLRLKGPREDHMD 179
 Db 121 TLHVLVTREYDLSAYEVDHGRGDYVSHIEMHQRERRRAVAVSEVESHLRLKGPREDHMD 180
 Qy 180 LRTSSSLVAPGFTVQTLGKTGTSVQTLPPEDFCFYQGSLSRHSRSPSGGKFCGSGTRT 239
 Db 181 LRTSSSLVAPGFTVQTLGKTGTSVQTLPPEDFCFYQGSLSRHSRSPSGGKFCGSGTRT 240
 Qy 240 LKLCNSOKCPDRSDVDFRAAQCABHNSRRFRGRHYKWKPYTQVEADLCKLYCIAEGDFDF 299
 Db 241 LKLCNSOKCPDRSDVDFRAAQCABHNSRRFRGRHYKWKPYTQVEADLCKLYCIAEGDFDF 300
 Qy 300 SLANKVKDGTFCSEDSRNVICIDICELSVVSTSAHMPQPKEDLFLPDEYKSCLRHRS 359
 Db 301 SLANKVKDGTFCSEDSRNVICIDICE-----NPQPKEDLFLPDEYKSCLRHRS 351
 Qy 360 LLRSRNEELNVETLVVDKMMQNHENITTYVLTILNMVSALFKQGLMKGKQTRHSH 419
 Db 352 LLRSRNEELNVETLVVDKMMQNHENITTYVLTILNMVSALFKQGLMKGKQTRHSH 403
 Qy 420 ALLTLGLDICSWKNEPCTLGFAPISGCMCKYRSCTINEDTGLGLAFTIAHSGHNGMI 479
 Db 404 ----- 403
 Qy 480 HDGEGNCKXSEGNISPTLAGRNGVFNPSPCSRQYLHKFLSTAQICLADQPKPVKEV 539
 Db 404 -----GNI-NIAIVG----- 412
 Qy 540 YPEKLPGELYDANTQCKWQFGEKAKLMDLDFKDKDICKALWCHRIKCKETKMPAAEGTI 599
 Db 413 -----LILLEDEQDICKALWCHRIKCKETKMPAAEGTI 447
 Qy 600 CGHDMWCRGQCVKYDGGPKPTHGNSDWSWSPCSRCCGGVSHRSRLCTNPKPSHCG 659
 Db 448 CGHDMWCRGQCVKYDGGPKPTHGNSDWSWSPCSRCCGGVSHRSRLCTNPKPSHCG 507
 Qy 660 KFCGSGTRTLKLCNSOKCPDRSDVDFRAAQCABHNSRRFRGRHYKWKPYTQVEDQDLCKLY 719
 Db 508 KFCGSGTRTLKLCNSOKCPDRSDVDFRAAQCABHNSRRFRGRHYKWKPYTQVEDQDLCKLY 561
 Qy 720 CIAEGDFPFLSLNKVKDGTFCSEDSRNVICIDICELSVVSTSAHMPQPKEDLFLPDEYKSCLRHRS 779
 Db 562 CIAEGDFPFLSLNKVKDGTFCSEDSRNVICIDICE--GCDNVLDGDAVEDVCGVCGN 619
 Qy 780 SACTIHRGLYTKHHHTNQYHYMTIPSGARSIRIYEMNVSTSYISVRNALRYLNGHWT 839
 Db 620 SACTIHRGLYTKHHHTN-YHYMTIPSGARSIRIYEMNVSTSYISVRNALRYLNGHWT 678
 Qy 840 VDWPGRYKFGSTTFDYSRYSNEPENLIATGPTNETLIVELLFQGRNPGVAVSEYMPRLGT 899
 Db 679 VDWPGRYKFGSTTFDYSRYSNEPENLIATGPTNETLIVELLFQGRNPGVAVSEYMPRLGT 738
 Qy 900 EKQPPAQPSYTWALVRSECVSGGGRCLPVLLEAACQPSATAYIALAFLES 952
 Db 739 EKQPPAQPSYTWALVRSECVSGGGRCLPVLLEAACQPSATAYIALAFLES 791

RESULT 4

AAU79497
 ID AAU79497 standard; Protein: 1224 AA.

XX AAU79497;

AC AAU79497;

DT 15-JUL-2002 (first entry)

XX Human ADAMTS protein #1.

DE Human ADAMTS protein #1.

XX Human; ADAMTS; cytoskeletal; antidiabetic; antirheumatic;

KW antiarthritic; antiulcer; vulnery; neovascularisation; angioma;

KW diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;

KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
 KW repolysin-type 2N-metalloprotease domain; disintegrin-like domain; TSPI;
 KW thrombospondin type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;
 XX chromosome 5p15.2-15.3; Cri-du-chat syndrome.
 OS Homo sapiens.
 XX W0200231163-A1.
 PN 18-APR-2002.
 XX 11-OCT-2001; 2001MO-JP08913.
 XX 11-OCT-2000; 2000JP-0311309.
 PR 02-APR-2001; 2001JP-0102905.
 XX (KAZU-) KAZUSA DNA RES INST FOUND.
 PA (MITS-) MITSUBISHI PHARMA CORP.
 XX Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 Kanazaki K;
 WIPI; 2002-372277/40.
 DR N-ESDB; ABK49822.
 XX Human brain-originated ADAMTS family polypeptide and encoded gene,
 PT applicable in diagnosis and screening compounds for drug compositions
 PT in treating diseases due to e.g. neovascularisation
 XX Claim 1; Page 146-150; 172pp; Japanese.
 XX The invention relates to a polypeptide belonging to the ADAMTS family is
 CC selected from sequences appearing as AAU79496, AAU79497 and AAU79499,
 CC a protein that contains the polypeptide, a protein having not less than
 CC 50% homology with the amino acid sequence of the polypeptides or a
 CC polypeptide modified from any of the polypeptides but with some amino
 CC acids deleted, substituted, added or inserted. Also included are the
 CC polynucleotides encoding the polypeptides (or their complementary
 CC strands or variants), a recombinant vector containing any of the
 CC polynucleotides, a transformant which is transformed with the recombinant
 CC vector, producing the polypeptide, protein or peptide by culturing the
 CC transformant, an antibody that can recognize the polypeptide, protein or
 CC peptide and screening compounds to promote or inhibit activity of the
 CC polypeptide or protein, or to promote or inhibit expression of the
 CC polynucleotide by using the polypeptide, protein, peptide, the
 CC polynucleotide, vector, transformant or/and antibody, particularly in
 CC the presence of a test compound for contact before evaluating the
 CC activity by measuring signal changes. The polypeptide and encoded gene
 CC are applicable in diagnosis and screening compounds for drug compositions
 CC in treating diseases due to neovascularisation, diabetic omentopathy,
 CC chronic rheumatoid arthritis, angioma, refractory skin and gastric ulcers
 CC and post-operative healing failure, including gene therapy.
 CC The gene encoding such polypeptide has conserved repolysin-type 2N-
 CC metalloprotease domain, disintegrin-like domain and TSPI (thrombospondin
 CC type 1) domain. Its encoded protein is characterised by high expression
 CC in ovaries, changes in expression dose depending on the sexual cycle, a
 CC decrease in tumour cell and location of the gene on the 5P-syndrome
 CC deletion site on chromosome 5p15.2-15.3 (associated with
 CC Cri-du-chat syndrome). The present sequence represents a ADAMTS
 CC protein of the invention.
 XX SQ Sequence 1224 AA;

Query Match 73.7%; Score 3858.5; DB 23; Length 1224;
 Best Local Similarity 75.6%; Pred. No. 3.1e-309;
 Matches 740; Conservative 24; Mismatches 72; Indels 143; Gaps 13;

Qy 1 MKPRARGWGLAALMLLAQVAVQSPGRSHORNGRSGOLEASPPRLSLRGPRLTAMS 60
 |||||
 Db 1 MKPRARGWGLAALMLLAQVAVQSPGRSHORNGRSGOLEASPPRLSLRGPRLTAMS 60
 |||||

Qy 61 PLFSAGTCVHRGTRSGSAWEPERPASSSTRGAGLDGKGRDMDGAGNHSQQTNGTENQ 120
 |||||
 |||||

Db 32 PAAAA-----PGSPVPRPPPAERAG-----WMEKG----- 58
 Qy 121 TLHLVLTQYDLVSAVEYDHRGCVSHIMHQRRAVAVSEVSHLRLKGRPHDFMDL 180
 Db 59 -----EYDLVLSAYEVDHGRDGVYSHIMHQRRAVAVSEVSHLRLKGRPHDFMDL 112
 Qy 181 RTSSSLVAPGFIVQTLGKTGTSVQTLPPEDFCFYQGSLSRSHNSPSHGKFCESGSTR 240
 Db 113 RTSSSLVAPGFIVQTLGKTGTSVQTLPPEDFCFYQGSLSRSHNS-SVALSTCOGLSGMI 171
 Qy 241 KLCNSQKCRPDSVDFRAAQAENHRRFRGR-----HYKWKPYTQVEA-DLCKLVCY 291
 Db 172 R-----TEEDYFLRPLPSHLSWL-GRAAQSGSPGHVLYKRSSTEPHAPGASEVLVT 222
 Qy 292 AEGFOFFFS--LGNKKVKGDFPCSEDSRNVCIIDGICELSVVSTSAHWPQPPKEDLFLPDE 349
 Db 223 SRTELAAHQPLHSDRLGLP---QKHFC-----GRKKYNPQPPKEDLFLPDE 270
 Qy 350 YKSLRKRSLRLSHRNEELNVETLVVVDKMMQNHENITTYVLTILNMYVALFKD-- 407
 Db 271 YKSLRKRSLRLSHRNEELNVETLVVVDKMMQNHENITTYVLTILNMYVALFKDGT 330
 Qy 408 -----GLMGKDGTRHDHAILLTGLD 427
 Db 331 IGININIAVGLILLDEQPLVISHADHTLSPQWQSLMGKDGTRHDHAILLTGLD 390
 Qy 428 ICSWKNEPCDTLGFARISGKSYRSCINEDTGLGLAFTIAHESGHNFMIHDEGNNC 487
 Db 391 ICSWKNEPCDTLGFARISGKSYRSCINEDTGLGLAFTIAHESGHNFMIHDEGNNC 450
 Qy 488 KSEGNIMSTLAGRNVFVSWSPCSROYLHKFLSTAQAICLADQPKVKEYKPEKLGE 547
 Db 451 KRSEGNIMSTLAGRNVFVSWSPCSROYLHKFLSTAQAICLADQPKVKEYKPEKLGE 510
 Qy 548 LYDANTOCKWQGEKAKLCLMDFKKDICKALWCHRIKRCETKFPAAEGTICGHDMMCR 607
 Db 511 LYDANTOCKWQGEKAKLCLMDFKKDICKALWCHRIKRCETKFPAAEGTICGHDMMCR 570
 Qy 608 GQCVKYGDGPPTHGHWSDSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 667
 Db 571 GQCVKYGDGPPTHGHWSDSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 630
 Qy 668 TLKLCNSQKCRPDSVDFRAAQAENHRRFRGRHYKWKPYTQVEDQDLCKLCIAEGDF 727
 Db 631 TLKLCNSQKCRPDSVDFRAAQAENHRRFRGRHYKWKPYTQVEDQDLCKLCIAEGDF 690
 Qy 728 FFLSNKVKDGTCCSEDSRNVCIIDGICERVCNDVLGSDAEDVCGVCGNNSACTIHRG 787
 Db 691 FFLSNKVKDGTCCSEDSRNVCIIDGICERVCNDVLGSDAEDVCGVCGNNSACTIHRG 750
 Qy 788 LYTKHHHTNQYHYHWTIPSGARSIRIYEMNVSTSYISVRNALRYLYLNGHWTVDWPGRYK 847
 Db 751 LYTKHHHTNQYHYHWTIPSGARSIRIYEMNVSTSYISVRNALRYLYLNGHWTVDWPGRYK 810
 Qy 848 FSGTTPDYRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAVSEYSHLRLKGRPHDFMDL 907
 Db 811 FSGTTPDYRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAVSEYSHLRLKGRPHDFMDL 870
 Qy 908 SYTWAIVRSECVSCGGGR 926
 Db 871 SYTWAIVRSECVSCGGGQ 889

RESULT 5

AAU72891

ID AAU72891 standard; Protein: 1224 AA.

AC AAU72891;

XX

DT 26-FEB-2002 (first entry)

XX

DE Human metalloprotease partial protein sequence #3.

XX

Human; protease; PCR primer; cytostatic; immunomodulator; cardiant; vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser; hypertensive; hypoglycaemic; neuroleptic; neuroprotective; anabolic; anorectic; antiinflammatory; aspartyl protease; cysteine protease; metalloprotease; serine protease; cancer; haematopoietic; breast; colon; lung; prostate; cervical; brain; ovarian; bladder; kidney; pain; immune-related disease; cardiovascular disease; neuronal disease; migraine; sexual dysfunction; mood disorder; attention disorder; cognition disorder; hypotension; hypertension; psychotic disorder; dyskinesia; metabolic disorder; inflammatory disorder.

Homo sapiens.

WO200183782-A2.

08-NOV-2001.

04-MAY-2001; 2001WO-US14431.

04-MAY-2000; 2000US-201879P.

(SUGB-) SUGEN INC.

Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S; Payne V;

WPI; 2002-041502/05.

N-PSDB; AAS97174.

Novel protease polypeptide useful for screening for substances that may be used to treat, e.g., cancers, immune-related diseases.

cardiovascular disease, migraine, pain, psychotic and inflammatory disorders -

Claim 28; Figure 2D; 232bp; English.

The invention relates to an isolated, enriched, or purified protease polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to screen for substances (S) that may modulate its activity. Administering S (which modulates protease activity in vitro) may be used to treat a disease or disorder selected from cancers (e.g., of tissues, of blood or haematopoietic origin, of the breast, colon, lung, prostate, cervical, brain, ovarian, bladder or kidney), immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases (e.g., central or peripheral nervous system diseases, migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders and dyskinesias), metabolic disorders and inflammatory disorders. (I) may also be useful as a diagnostic tool for a disease or disorder such as those above. AAU72876-AAU72910 represent human protease amino acid sequences of the invention.

Sequence 1224 AA;

Query Match 73.7%; Score 3859.5; DB 23; Length 1224;

Best Local Similarity 75.6%; Pred. No. 3.1e-309;

Matches 740; Conservative 24; Mismatches 72; Indels 143; Gaps 13;

Qy 1 MKPRARGWGLAALWMLLAQVAVQSPGRSHQGRNGSGLEASPRLLSRGRPRUTAMS 60

Db 1 MKPRARGWGLAALWMLLAQVAVQSPGRSHQGRNGSGLEASPRLLSRGRPRUTAMS 31

Qy 61 PLFSAGTCVRHGTGTSAMEPERPASSTRGAAGLDGKGRDMDAGNHRSSQNTGTENQ 120

Db 32 PAAAA-----PGSPVPRPPPAERAG-----WMEKG----- 58

Qy 121 TLHLVLTQYDLVSAVEYDHRGCVSHIMHQRRAVAVSEVSHLRLKGRPHDFMDL 180

Db 59 -----EYDLVLSAYEVDHGRDGVYSHIMHQRRAVAVSEVSHLRLKGRPHDFMDL 112

Qy 181 RTSSSLVAPGFIVQTLGKTGTSVQTLPPEDFCFYQGSLSRSHNSPSHGKFCESGSTR 240

Db 113 RTSSSLVAPGFIVQTLGKTGTSVQTLPPEDFCFYQGSLSRSHNS-SVALSTCOGLSGMI 171

QY 241 KLCNSQKPRDSVDFRAAQAENRRFRGR-----HYKWKPYTQVEA-DLCKLYCI 291
 DB 172 R-----TEADYFLRPLPUSHLWKL-GRAAQSSPSHVLYKRSTEPHAPGASEVLVT 222
 QY 292 AEGFDFFFS--LSNKVKDGTGTCSEDSRNVICIDIGICELSVSVSAHNPQPKEDLFLPDE 349
 DB 223 SRTWELAQPLHSSDLRLGLP---QKQHF-----GRRKKYMPQPKEDLFLPDE 270
 QY 350 YKSLRHRKRLSRHRNEELNVTLVVVDKMMQNHGHEHNTTIVTLILNWSALFKD-- 407
 DB 271 YKSLRHRKRLSRHRNEELNVTLVVVDKMMQNHGHEHNTTIVTLILNWSALFKDGT 330
 QY 408 -----GLMKDGTTRHDHAILLTGLD 427
 DB 331 IGGNINIAVGLILLEDEQPLGVISHHADHTLSSFCQWQSGLMGKDGTRHDHAILLTGLD 390
 QY 428 ICSWKNPCDPTLGFAPISGCMKYSRSTCTINEDTGLGLAFTIAHESGHNFCMTHDGSNMC 487
 DB 391 ICSWKNPCDPTLGFAPISGCMKYSRSTCTINEDTGLGLAFTIAHESGHNFCMTHDGSNMC 450
 QY 488 KKEGNIPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPVKEYKPEKLPGE 547
 DB 451 KKEGNIPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPVKEYKPEKLPGE 510
 QY 548 LYDANTOCKWQFGKAKLMLDFKKDICKALWCHRIKRCETKFMFAAESTICGHDMWCR 607
 DB 511 LYDANTOCKWQFGKAKLMLDFKKDICKALWCHRIKRCETKFMFAAESTICGHDMWCR 570
 QY 608 GGQCVKYGDEGPKPTHGHWSDWSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCBGSTR 667
 DB 571 GGQCVKYGDEGPKPTHGHWSDWSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCBGSTR 630
 QY 668 TLKLCNSQKPRDSVDFRAAQAENRRFRGRHYKWKPYTQVEDDOLCKLYCIAEGFDF 727
 DB 631 TLKLCNSQKPRDSVDFRAAQAENRRFRGRHYKWKPYTQVEDDOLCKLYCIAEGFDF 690
 QY 728 FFSLSNKVKDGTGTCSEDSRNVICIDIGICERYGCDNLGSDAIVEDVCGVCGNNSACTIHRG 787
 DB 691 FFSLSNKVKDGTGTCSEDSRNVICIDIGICERYGCDNLGSDAIVEDVCGVCGNNSACTIHRG 750
 QY 788 LYTKHHHTNQYHHMTTPSGARSIRIYEMNVSTYSISVRNALRYVILNGHMTVDWPGRYK 847
 DB 751 LYTKHHHTNQYHHMTTPSGARSIRIYEMNVSTYSISVRNALRYVILNGHMTVDWPGRYK 810
 QY 848 PSGTTFDYRRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQ 907
 DB 811 PSGTTFDYRRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQ 870
 QY 908 SYTWAIVRSECVSCGGGR 926
 DB 871 SYTWAIVRSECVSCGGGQ 889

RESULT 6

AAO15254
 ID AAO15254 standard; Protein; 1224 AA.

XX AAO15254;

XX AAO15254;

DT 05-SEP-2002. (first entry)

DE Human MDT99 protease amino acid sequence.

XX Human; MDT99 protease; TGF-beta inhibitor;
 KW transforming growth factor-beta inhibitor; chronic renal failure.
 XX Homo sapiens.

XX WO200251998-A1.

XX 04-JUL-2002.

XX

PF 21-DEC-2001; 2001WO-JP11251.
 XX 25-DEC-2000; 2000JP-0393372.
 XX (YAMA) YAMANOUCHI PHARM CO LTD.
 XX Yamaji N, Nishimura K, Abe K, Ogino M;
 XX WPI; 2002-50888/54.
 XX N-PSDB; AAL43654.
 PT Protease MDT99 and encoded polynucleotide, applicable in diagnosis and
 PT screening TGF-beta inhibitors for treatment of chronic renal failure
 XX Claim 1; Page 48-51; 60pp; Japanese.
 CC The invention comprises the amino acid and coding sequence of the human
 CC MDT99 protease. The MDT99 DNA and protein sequences of the invention are
 CC useful for screening for transforming growth factor (TGF)-beta inhibitors
 CC and for the treatment of chronic renal failure. The present amino acid
 CC sequence represents the human MDT99 protease of the invention.
 XX Sequence 1224 AA;
 SQ
 Query Match 73.5%; Score 3846.5; DB 23; Length 1224;
 Best Local Similarity 75.4%; Pred. No. 3e-308;
 Matches 738; Conservative 25; Mismatches 73; Indels 143; Gaps 13;
 QY 1 MKPRARGWGLAALMLLAQVAEQVSPGRSHQRNGRSGQLEASPPRLSRGPRRLTAMS 60
 DB 1 MKPRARGWGLAALMLLAQVAEQAP-----ACAMG 31
 QY 61 PLFAGTCTVRHGTSSGSAWEPEPSSSTRGAAGLDGKGRDMDGAGNHRSQNTGTENQ 120
 DB 32 PAAIA-----PGSPVPRPPAERPQ-----WMEKG----- 58
 QY 121 TLHVLTVQDLSAYEVDHGRDYSHEIMHQRRAVAVSEVSLHLRLKGRPHDFHMDL 180
 DB 59 -----EYDLVSAVEVDHGRDYSHEIMHQRRAVAVSEVSLHLRLKGRPHDFHMDL 112
 QY 191 RTSSSLNAPGFIIVTGLTKTSVOTLPPEDFCFYQSLRSHRNSPSHGKFCGSTR 240
 DB 113 RTSSSLNAPGFIIVTGLTKTSVOTLPPEDFCFYQSLRSHRNS-SVALSTCGLSGMI 171
 QY 241 KLCNSQKPRDSVDFRAAQAENRRFRGR-----HYKWKPYTQVEA-DLCKLYCI 291
 DB 172 R-----TEADYFLRPLPUSHLWKL-GRAAQSSPSHVLYKRSTEPHAPGASEVLVT 222
 QY 292 AEGFDFFFS--LSNKVKDGTGTCSEDSRNVICIDIGICELSVSVSAHNPQPKEDLFLPDE 349
 DB 223 SRTWELAQPLHSSDLRLGLP---QKQHF-----GRRKKYMPQPKEDLFLPDE 270
 QY 350 YKSLRHRKRLSRHRNEELNVTLVVVDKMMQNHGHEHNTTIVTLILNWSALFKD-- 407
 DB 271 YKSLRHRKRLSRHRNEELNVTLVVVDKMMQNHGHEHNTTIVTLILNWSALFKDGT 330
 QY 408 -----GLMKDGTTRHDHAILLTGLD 427
 DB 331 IGGNINIAVGLILLEDEQPLGVISHHADHTLSSFCQWQSGLMGKDGTRHDHAILLTGLD 390
 QY 428 ICSWKNPCDPTLGFAPISGCMKYSRSTCTINEDTGLGLAFTIAHESGHNFCMTHDGSNMC 487
 DB 391 ICSWKNPCDPTLGFAPISGCMKYSRSTCTINEDTGLGLAFTIAHESGHNFCMTHDGSNMC 450
 QY 488 KKEGNIPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPVKEYKPEKLPGE 547
 DB 451 KKEGNIPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPVKEYKPEKLPGE 510
 QY 548 LYDANTOCKWQFGKAKLMLDFKKDICKALWCHRIKRCETKFMFAAESTICGHDMWCR 607
 DB 511 LYDANTOCKWQFGKAKLMLDFKKDICKALWCHRIKRCETKFMFAAESTICGHDMWCR 570
 QY 608 GGQCVKYGDEGPKPTHGHWSDWSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCBGSTR 667

Db 571 GGQCVKYGDEGPKPTHGHSWMSWSPCSRCTCGGVSHRSLCTNPKPSHGKFCBGSTR 630
 Qy 668 TLKLCNSQKCPDRSDVDFRAAQAQCAEHSRRFRGRHYKWKPYTQVEDQDLCKLYCIAEGFDF 727
 Db 631 TLKLCNSQKCPDRSDVDFRAAQAQCAEHSRRFRGRHYKWKPYTQVEDQDLCKLYCIAEGFDF 690
 Qy 728 FFSLSNKKVKGDTPCSEDSNRNVCIDGICERVGCNDVLGSDADEVCGVCGNNGNSACTIHRG 787
 Db 691 FFSLSNKKVKGDTPCSEDSNRNVCIDGICERVGCNDVLGSDADEVCGVCGNNGNSACTIHRG 750
 Qy 788 LYTKHHHTNQYHHWVTIPSGARSIRIYEMNVSTSYISVRNALRRYYLNGHWTVDWPGRYK 847
 Db 751 LYTKHHHTNQYHHWVTIPSGARSIRIYEMNVSTSYISVRNALRRYYLNGHWTVDWPGRYK 810
 Qy 848 FSGTTDFYRRSYNPNENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQP 907
 Db 811 FSGTTDFYRRSYNPNENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQP 870
 Qy 908 SYTWAIVRSECSVSCGGGR 926
 Db 871 SYTWAIVRSECSVSCGGGQ 889

RESULT 7

ABG72431 standard; Protein; 1224 AA.

XX AC ABG72431;

XX DT 05-FEB-2003 (first entry)

XX DE Novel human protease #2.

XX KW Human; protease; gene therapy; obesity; enzyme.

XX OS Homo sapiens.

XX PN US6448388-B1.

XX PD 10-SEP-2002.

XX PF 15-AUG-2001; 2001US-0930872.

XX PR 16-AUG-2000; 2000US-225852P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Friddle CU, Hilbun E;

XX DR WPI; 2003-074103/07.

XX DR N-PSDB; ABS57767, ABS57768.

PT New nucleic acid encoding novel human protein (NHP), useful for the preparation of a medicament for diagnosing, preventing or treating disorders or diseases associated with the NHP, e.g., obesity -

PS Claim 4; Column 19-26; 17pp; English.

CC The invention describes a new isolated nucleic acid comprising a sequence that encodes a fully defined protein sequence comprising 491 or 1224 amino acids or that hybridises under stringent conditions with the 3675-bp sequence or its complement. The nucleic acid is useful for the preparation of a medicament for diagnosing, preventing or treating disorders (e.g. using gene therapy) or diseases associated with the novel human proteins, e.g., obesity. This sequence encodes a novel human protease.

XX Sequence 1224 AA;

Query Match 73.5%; Score 3846.5; DB 24; Length 1224;

Best Local Similarity 75.4%; Pred. No. 3e-308;

Matches 738; Conservative 25; Mismatches 73; Indels 143; Gaps 13;

Qy 1 MKPRARGWRLAALMLLAQVAEQVSPGRSHQRNGRSGOLEASPPRLLSRGRRLUTAMS 60
 Db 1 MKPRARGWRLAALMLLAQVAEQVSPGRSHQRNGRSGOLEASPPRLLSRGRRLUTAMS 60
 Qy 61 PLFSAGTCVPHGTRSSGAMPEPASPSTSTGAGLOCKGDMDEAGNHRSQQTNTGTENQ 120
 Db 32 PAAAA-----PUSPSVPRPPPAERPG-----WNEKG-----58
 Qy 121 TLHLVLTQYDLVSAYEDHRODYVSHETIMHQRRAVAVSEVESLHLRLKGRPHDMOL 180
 Db 59 -----EYDLVSAYEDHRODYVSHETIMHQRRAVAVSEVESLHLRLKGRPHDMOL 112
 Qy 181 RTSSSLVAPGFIYQTLGKTGTSVQTLPPEDFCFYQGSLSHRNSPSHGKFCGEGSTRTL 240
 Db 113 RTSSSLVAPGFIYQTLGKTGTSVQTLPPEDFCFYQGSLSHRNS-SVALSTCGSLGMI 171
 Qy 241 KLCNSQKCPDRSDVDFRAAQAQCAEHSRRFRGR-----HYKWKPYTQVEA-DLCKLYCI 291
 Db 172 R-----TEADYFLRPLPSHLSWKL-GRAAQGSPPSHVLYKRSTEPHAPGASEVLVT 222
 Qy 292 AEGDFDFFS--LSNKKVKGDTPCSEDSNRNVCIDGICELSVVSTSAHMPQPPKEDLFIPLDE 349
 Db 223 SRTWELAHQPLHSSDLRLGLP---QKQHF-----GRRKKYMPQPPKEDLFIPLDE 270
 Qy 350 YKSLRHKRSLRSHRNEELNVETLVVDKMKMNHGHNITTYVLTILNMVSALSKD--407
 Db 271 YKSLRHKRSLRSHRNEELNVETLVVDKMKMNHGHNITTYVLTILNMVSALSKDGT 330
 Qy 408 -----GLMGKDGTRHDHAILLTGLD 427
 Db 331 IGGNINIAVGLILLEDEQPLVISHADHTLSSFCWQSGGLMGKDGTRHDHAILLTGLD 390
 Qy 428 ICSWKNEPCDTLGFAPISGMCYSKYRSTINEDTGLGLAFTIAHESGHNFGMIHDGEGNMC 487
 Db 391 ICSWKNEPCDTLGFAPISGMCYSKYRSTINEDTGLGLAFTIAHESGHNFGMIHDGEGNMC 450
 Qy 488 KKSEGNIMSPTLAGRNQVFSWSPCSRYLHKFLSTAQAICLADQPKVPKYEKYPEKLPG 547
 Db 451 KKSEGNIMSPTLAGRNQVFSWSPCSRYLHKFLSTAQAICLADQPKVPKYEKYPEKLPG 510
 Qy 548 LYDANTQCKWQFGEKAKLWLDFFKDI CKALWCHRI GRKCEKTFMPAAEGTICGHDMWCR 607
 Db 511 LYDANTQCKWQFGEKAKLWLDFFKDI CKALWCHRI GRKCEKTFMPAAEGTICGHDMWCR 570
 Qy 608 GGQCVKYGDEGPKPTHGHSWMSWSPCSRCTCGGVSHRSLCTNPKPSHGKFCBGSTR 667
 Db 571 GGQCVKYGDEGPKPTHGHSWMSWSPCSRCTCGGVSHRSLCTNPKPSHGKFCBGSTR 630
 Qy 668 TLKLCNSQKCPDRSDVDFRAAQAQCAEHSRRFRGRHYKWKPYTQVEDQDLCKLYCIAEGFDF 727
 Db 631 TLKLCNSQKCPDRSDVDFRAAQAQCAEHSRRFRGRHYKWKPYTQVEDQDLCKLYCIAEGFDF 690
 Qy 728 FFSLSNKKVKGDTPCSEDSNRNVCIDGICERVGCNDVLGSDADEVCGVCGNNGNSACTIHRG 787
 Db 691 FFSLSNKKVKGDTPCSEDSNRNVCIDGICERVGCNDVLGSDADEVCGVCGNNGNSACTIHRG 750
 Qy 788 LYTKHHHTNQYHHWVTIPSGARSIRIYEMNVSTSYISVRNALRRYYLNGHWTVDWPGRYK 847
 Db 751 LYTKHHHTNQYHHWVTIPSGARSIRIYEMNVSTSYISVRNALRRYYLNGHWTVDWPGRYK 810
 Qy 848 FSGTTDFYRRSYNPNENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQP 907
 Db 811 FSGTTDFYRRSYNPNENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQP 870
 Qy 908 SYTWAIVRSECSVSCGGGR 926
 Db 871 SYTWAIVRSECSVSCGGGQ 889

RESULT 8

AAU80153

ID AAU80153 standard; Protein; 1224 AA.

XX AC AAU80153;
 XX DT 15-JUL-2002 (first entry)
 XX DE Human ADAMTS protein variant.
 XX DE
 XX KW Human; ADAMTS; cytostatic; antidiabetic; antirheumatic;
 KW antiarthritic; antitumor; vulnery; neovascularisation; angioma;
 KW diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
 KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
 KW repolysin-type 2N-metalloproteinase domain; disintegrin-like domain; TSPI;
 KW thrombospondin type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;
 KW chromosome 5p15.2-15.3; Cri-du-chat syndrome.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Misc-difference 104
 FT /note= "Wild-type Pro substituted by Ser"
 FT Misc-difference 110
 FT /note= "Wild-type Met substituted by Val"
 FT Misc-difference 284
 FT /note= "Wild-type Ser substituted by Thr"
 XX WO200231163-A1.
 XX PN 18-APR-2002.
 XX PD
 XX 11-OCT-2001; 2001WO-JP08913.
 XX PF 11-OCT-2000; 2000JP-0311309.
 XX PR 02-APR-2001; 2001JP-0102905.
 XX PX (KAZU-) KAZUSA DNA RES INST FOUND.
 PX (MITS-) MITSUBISHI PHARMA CORP.
 XX PI Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 PI Kanzaki K;
 XX WPI; 2002-372277/40.
 XX Human brain-originated ADAMTS family polypeptide and encoded gene,
 FT applicable in diagnosis and screening compounds for drug compositions
 FT in treating diseases due to e.g. neovascularisation.
 XX PS Example 2; Page -: 172pp; Japanese.
 XX The invention relates to a polypeptide belonging to the ADAMTS family is
 CC selected from sequences appearing as AAU79496, AAU79497 and AAU79499,
 CC a protein that contains the polypeptide, a protein having not less than
 CC 50% homology with the amino acid sequence of the polypeptides or a
 CC polypeptide modified from any of the polypeptides but with some amino
 CC acids deleted, substituted, added or inserted. Also included are the
 CC polynucleotides encoding the polypeptides (or their complementary
 CC strands or variants), a recombinant vector containing any of the
 CC polynucleotides, a transformant which is transformed with the recombinant
 CC vector, producing the polypeptide, protein or peptide by culturing the
 CC transformant, an antibody that can recognize the polypeptide, protein or
 CC peptide and screening compounds to promote or inhibit activity of the
 CC polypeptide or protein, or to promote or inhibit expression of the
 CC polynucleotide by using the polypeptide, protein, peptide, the
 CC polynucleotide, vector, transformant or/and antibody, particularly in
 CC the presence of a test compound for contact before evaluating the
 CC activity by measuring signal changes. The polypeptide and encoded gene
 CC are applicable in diagnosis and screening compounds for drug compositions
 CC in treating diseases due to neovascularisation, diabetic omentopathy,
 CC chronic rheumatoid arthritis, angioma, refractory skin and gastric ulcers
 CC and post-operative healing failure, including gene therapy.
 CC The gene encoding such polypeptide has conserved repolysin-type 2N-
 CC metalloproteinase domain, disintegrin-like domain and TSPI (thrombospondin
 CC type 1) domain. Its encoded protein is characterised by high expression
 CC in ovaries, changes in expression dose depending on the sexual cycle, a

CC decrease in tumour cell and location of the gene on the 5P-syndrome
 CC deletion site on chromosome 5p15.2-15.3 (associated with
 CC Cri-du-chat syndrome). The present sequence represents a variant of
 CC the human ADAMTS protein of the invention.
 XX
 XX SQ Sequence 1224 AA;
 Query Match 73.3%; Score 3840.5; DB 23; Length 1224;
 Best Local Similarity 75.3%; Pred. No. 9.5e-308;
 Matches 737; Conservative 25; Mismatches 74; Indels 143; Gaps 13;
 QY 1 MKPARGWGLAALMLLAQVAEQVSPGRSHQGRNGSGQLEASPPRLISRGPRRLTAMS 60
 DB 1 MKPARGWGLAALMLLAQVAEQAP-----ACAMG 31
 QY 61 PLFSAGTCVRHGTSGSAMPEFPAASSTGGAAGLCKGKDRMDDEAGNHRSSQNTGTENQ 120
 DB 32 PAAAA-----PGSFVSFPRPPPPAERF-----MMKEG----- 58
 QY 121 TLHYLTQYDLVSAYEVDHGDYSHZIMHHQRRRAVAVSEVSLHLRLKGRPHDMDL 180
 DB 59 -----BYDLVSAYEVDHGDYSHZIMHHQRRRAVAVSEVSLHLRLKGRPHDMDL 112
 QY 181 RTSSSLVAPGFIVQTLGKTCTKSYQTLPPEDFCFYQGSLSRSHNSPFCGSGSTRTL 240
 DB 113 RTSSSLVAPGFIVQTLGKTCTKSYQTLPPEDFCFYQGSLSRSHNS-SVALSTCQGLSGMI 171
 QY 241 KLCNSQKCPRDSVDFRAAQCAEHSRRFRGR-----HYKWKPTQVFA-DLCKLYCI 291
 DB 172 R-----TEADYFLRPLPSHLWSKL-GRAAQGSPSHVLYKRSSTEPHAPASEVLVT 222
 QY 292 AEGDFPFFS--LSNKVKGDTGFCSEDSRNVCIDGICELSVVSTSAHMPQPPKEDFILPDE 349
 DB 223 SRTWELAHQPLHSSDLRLGLP---QKQHF-----GRRKKYMPQPKEDFILPDE 270
 QY 350 YKSCLRHKLRLSRHNSLNVELVVDKXKQNHGHENITTYVLTILNWSALFKD-- 407
 DB 271 YKSCLRHKLRLRYRNEELNVELVVDKXKQNHGHENITTYVLTILNWSALFKDGT 330
 QY 408 -----GLMKQDQTRHDHAILLTGLD 427
 DB 331 IGGNITAIIVGLILLEDEQPLVISHHADHTLSFQWQSGLMGKQDTRHDHAILLTGLD 390
 QY 428 ICSWKNPDCDTLGPAPISGMSKYRSCITNEDTGLGLAFTIAHESGHNFMHIDGEGNMC 487
 DB 391 ICSWKNPDCDTLGPAPISGMSKYRSCITNEDTGLGLAFTIAHESGHNFMHIDGEGNMC 450
 QY 488 KKSNGTMSPLAGNCFVSWSPCSROYLHKLSTQAACLADQPKVPKVKYKPEKLPGE 547
 DB 451 KKSNGTMSPLAGNCFVSWSPCSROYLHKLSTQAACLADQPKVPKVKYKPEKLPGE 510
 QY 548 LYDANTQCKWQFGEKAKLCLMDPKKDTCKALWCHRIKRCETKFMFAAEGTICGHDMWCR 607
 DB 511 LYDANTQCKWQFGEKAKLCLMDPKKDTCKALWCHRIKRCETKFMFAAEGTICGHDMWCR 570
 QY 608 GGQCVKYGDEGPKPTHGHSWDSWSPCSRTCGGSHSRSELCTNPKPSGGHFCGSGSTR 667
 DB 571 GGQCVKYGDEGPKPTHGHSWDSWSPCSRTCGGSHSRSELCTNPKPSGGHFCGSGSTR 630
 QY 668 TLKLCNSQKCPRDSVDFRAAQCAEHSRRFRGRHYKWKPTQVDEDQDLCKLYCIAEGFDF 727
 DB 631 TLKLCNSQKCPRDSVDFRAAQCAEHSRRFRGRHYKWKPTQVDEDQDLCKLYCIAEGFDF 690
 QY 728 PFSLSNKVKDGTGTPCSEDSRNVCIDGICERVCNDVIGSDAVEDVCGVCGNNSACTIHRG 787
 DB 691 PFSLSNKVKDGTGTPCSEDSRNVCIDGICERVCNDVIGSDAVEDVCGVCGNNSACTIHRG 750
 QY 788 LYTGHHTNYHYHMTVTPSGARSTRITYEMNVSTSYLSVRNALRYLYNGHWTVDWPGRYK 847
 DB 751 LYTGHHTNYHYHMTVTPSGARSTRITYEMNVSTSYLSVRNALRYLYNGHWTVDWPGRYK 810
 QY 848 FSGTTFDYRRSYNBPENLIATGNETPLIVELLFQGRNFCGVAWEYSMPRLGTGKQPPAOP 907

Db 811 FSGTTDFRRSYNEPENLIGTNETLIVELLFQGRNPGVAVWEYMPRLGTEKQPPAQ 870

QY 908 SYTVAIVRSECVSCGGQ 926

Db 871 SYTVAIVRSECVSCGGQ 889

RESULT 9

AAU79500
ID AAU79500 standard; Protein; 1170 AA.

XX AC AAU79500;

XX 15-JUL-2002 (first entry)

XX Human ADAMTS protein #3.

XX Human; ADAMTS; cytostatic; antidiabetic; antirheumatic;
KW antiarthritic; antiulcer; vulnary; neovascularisation; angioma;
KW diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
KW reopolyisin-type 2N-metalloprotease domain; disintegrin-like domain; TSP1;
KW reopolyisin-type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;
KW chromosome 5p15.2-15.3; Cri-du-chat syndrome.

XX Homo sapiens.

XX W0200231163-A1.

XX 18-APR-2002.

XX 11-OCT-2001; 2001WO-JP08913.

XX 11-OCT-2000; 2000JP-0311309.

XX 02-APR-2001; 2001JP-0102905.

XX (KAZU-) KAZUSA DNA RES INST FOUND.

XX (MITS-) MITSUBISHI PHARMA CORP.

XX Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;

XX Kanzaki K;

XX WPI; 2002-372277/40.

XX Human brain-originated ADAMTS family polypeptide and encoded gene,
PT applicable in diagnosis and screening compounds for drug compositions
PT in treating diseases due to e.g. neovascularisation

XX Example 8; Fig 8; 172pp; Japanese.

XX The invention relates to a polypeptide belonging to the ADAMTS family is
XX selected from sequences appearing as AAU79495, AAU79497 and AAU79499,
XX a protein that contains the polypeptide, a protein having not less than
XX 50% homology with the amino acid sequence of the polypeptides or a
XX polypeptide modified from any of the polypeptides but with some amino
XX acids deleted, substituted, added or inserted. Also included are the
XX polynucleotides encoding the polypeptides (or their complementary
XX strands or variants), a recombinant vector containing any of the
XX polynucleotides, a transformant which is transformed with the recombinant
XX vector, producing the polypeptide, protein or peptide by culturing the
XX transformant, an antibody that can recognize the polypeptide, protein or
XX peptide and screening compounds to promote or inhibit activity of the
XX polypeptide or protein, or to promote or inhibit expression of the
XX polynucleotide by using the polypeptide, protein, peptide, the
XX polynucleotide, vector, transformant or/and antibody, particularly in
XX the presence of a test compound for contact before evaluating the
XX activity by measuring signal changes. The polypeptide and encoded gene
XX are applicable in diagnosis and screening compounds for drug compositions
XX in treating diseases due to neovascularisation, diabetic omentopathy,
XX chronic rheumatoid arthritis, angioma, refractory skin and gastric ulcers
XX and post-operative healing failure, including gene therapy.
XX The gene encoding such polypeptide has conserved reopolyisin-type 2N-
XX metalloprotease domain, disintegrin-like domain and TSP1 (thrombospondin

CC type 1) domain. Its encoded protein is characterised by high expression
CC in ovaries, changes in expression dose depending on the sexual cycle, a
CC decrease in tumour cell and location of the gene on the 5P-syndrome
CC deletion site on chromosome 5p15.2-15.3 (associated with
CC Cri-du-chat syndrome). The present sequence represents a ADAMTS
CC protein of the invention.

XX SQ Sequence 1170 AA;

Query Match 71.6%; Score 3750.5; DB 23; Length 1170;
Best Local Similarity 82.3%; Pred. No. 2.4e-300;
Matches 702; Conservative 23; Mismatches 53; Indels 75; Gaps 9;

QY 127 QYDLVSAYEVDHREDYVSHHEIMHQRRAVAVSEVESLHLRLKGRPHDRLTSSSL 186
Db 5 EYDLVSAYEVDHREDYVSHHEIMHQRRAVAVSEVESLHLRLKGRPHDRLTSSSL 64

QY 187 VAFGFIQTLGKTGTSVOTLPPEDCFYQGSLSRHSRSPSHGGKFCGEGSTTLKLCNQ 246

Db 65 VAFGFIQTLGKTGTSVOTLPPEDCFYQGSLSRHSRNS-SVALSTCQGLSMIR----- 118

QY 247 KCPDRSDVDFRAACCAHNGRRFRGR-----HYWKKEPYTQVEA-DLCKLYCIAEGFDF 297

Db 119 ---TEADYFLRPLPSHLNWKL-GRAAQGSFPHVLYKSTPHAPGASEVLVTSRTWEL 174

QY 298 FFS--LSNVKVGDTGTCSEDSNRNVCIDGICELSVVTSAMHPPQPPKEDLFIIDPKYKSLR 355

Db 175 AHQPLASDRLGLP---QKQFEC-----GREKKYMPQPPKEDLFIIDPKYKSLR 222

QY 356 HKESLLSRHNEELNVETLVVDKQWQNHENITTYVLTILNMVSAALFKO----- 407

Db 223 HKESLLSRHNEELNVETLVVDKQWQNHENITTYVLTILNMVSAALFKO----- 282

QY 408 -----GLMGKDGTRHDHAILLTGLDICSWN 433

Db 283 IATVGLILLEDEQGLVISHHADHTLSSFCQMSQGLMGKDGTRHDHAILLTGLDICSWN 342

QY 434 EPDCTLGFAPIGKSKYRSCCTINEDTGLAETIAHESGHNFGMTHDGNNCKSEGN 493

Db 343 EPDCTLGFAPIGKSKYRSCCTINEDTGLAETIAHESGHNFGMTHDGNNCKSEGN 402

QY 494 IMSEPLAGRNQVFSWSPCSRQYLHKLFLSTAQAICLADQPKPKYKPKYKPEKLPGLYDANT 553

Db 403 IMSEPLAGRNQVFSWSPCSRQYLHKLFLSTAQAICLADQPKPKYKPKYKPEKLPGLYDANT 462

QY 554 QCKWQFGEKAKLQMLDFKDKICKALMCHRIKRCETKEMPAARGTICGHDMWCRGGQCVK 613

Db 463 QCKWQFGEKAKLQMLDFKDKICKALMCHRIKRCETKEMPAARGTICGHDMWCRGGQCVK 522

QY 614 YGDEGPKPTHGHWSDSWSPCSRQYLHKLFLSTAQAICLADQPKPKYKPKYKPEKLPGLYDANT 673

Db 523 YGDEGPKPTHGHWSDSWSPCSRQYLHKLFLSTAQAICLADQPKPKYKPKYKPEKLPGLYDANT 582

QY 674 SQKCPRSDVDFRAACCAHNSRFRGRHRYKWKPYTQVEODDLCKLYCIAEGFDFFFLSLN 733

Db 583 SQKCPRSDVDFRAACCAHNSRFRGRHRYKWKPYTQVEODDLCKLYCIAEGFDFFFLSLN 642

QY 734 KVKGDTGTCSEDSNRNVCIDGICERVGCDNVLGSDAEDVCGVCGNNGNSACTIHRGLYTKH 793

Db 643 KVKGDTGTCSEDSNRNVCIDGICERVGCDNVLGSDAEDVCGVCGNNGNSACTIHRGLYTKH 702

QY 794 HTNYYHMVTPSGARSIRIYENNVSTSYISVNALRRYLYNGHWTVDPGRYKPSGTTTF 853

Db 703 HTNYYHMVTPSGARSIRIYENNVSTSYISVNALRRYLYNGHWTVDPGRYKPSGTTTF 762

QY 854 DYRSYNEPENLIGTNETLIVELLFQGRNPGVAVWEYMPRLGTEKQPPAQPSYTWAI 913

Db 763 DYRSYNEPENLIGTNETLIVELLFQGRNPGVAVWEYMPRLGTEKQPPAQPSYTWAI 822

QY 914 VRSECVSCGGQ 926

Db 823 VRSECVSCGGQ 835

RESULT 10	XX	PS	Claim 1; Page 18; 305pp; English.
ABG76896	XX	CC	The present invention relates to new NOVX (NOV1-10) polypeptides. The
ID ABG76896 standard; Protein; 856 AA.	XX	CC	molecules of the invention are useful for treating or preventing a
XX	XX	CC	NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or
AC ABG76896;	XX	CC	a disorder related to cell signal processing and metabolic pathway,
XX	XX	CC	modulation in humans. NOVX polypeptides, nucleic acids and antibodies
DT 05-NOV-2002 (first entry)	XX	CC	are useful for treating or preventing disorders or syndromes including
XX	XX	CC	breast cancer, Alzheimer's disease, epilepsy, Huntington's disease,
DE Human ADAM-TS 7-like protein #1.	XX	CC	anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis,
XX	XX	CC	neurodegeneration, Parkinson's disease, pain, stroke, autoimmune
XX	XX	CC	disease, allergies, addiction, asthma, endometriosis, graft versus host
KW Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing;	XX	CC	disease, systemic lupus erythematosus, scleroderma, transplantation,
KW breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;	XX	CC	psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,
KW anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;	XX	CC	atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes,
KW neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;	XX	CC	autoimmune disease; allergy; addiction; asthma; transplantation;
KW graft versus host disease; systemic lupus erythematosus; scleroderma;	XX	CC	thrombocytopenia; bleeding disorders, metabolic disorders, obesity,
KW psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;	XX	CC	atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
KW atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;	XX	CC	thrombocytopenia; bleeding disorders; glomerulonephritis; hypercalcaemia;
KW glucose transport defect; glomerulonephritis; hypercalcaemia;	XX	CC	polycystic kidney disease; renal tubular acidosis; skin disorder;
KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;	XX	CC	muscle disorder; bone disorder; joint disorder; skeletal disorder;
KW haematopoietic disorder; urinary system disorder; osteoporosis;	XX	CC	dental disease; dental infection; growth disorder; reproductive disorder;
KW hypogonadism; fertility disorder; viral infection; bacterial infection;	XX	CC	parasitic infection; metabolic pathway modulation; gene therapy;
KW zinc metallopeptidase; ADAM-TS 7; alpha-2-macroglobulin precursor;	XX	CC	ileal sodium/bile acid cotransporter; prohibitin; MT; CIP4; spinosin;
KW macrophage stimulating protein precursor; fatty acid-binding protein;	XX	CC	gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.
XX	OS	Homo sapiens.	
XX	WO200233087-A2.		
XX	25-APR-2002.		
XX	17-OCT-2001; 2001WO-US32496.		
XX	17-OCT-2000; 2000US-241040P.		
PR	17-OCT-2000; 2000US-241058P.		
PR	17-OCT-2000; 2000US-241063P.		
PR	17-OCT-2000; 2000US-241243P.		
PR	20-OCT-2000; 2000US-242152P.		
PR	23-OCT-2000; 2000US-242482P.		
PR	23-OCT-2000; 2000US-242611P.		
PR	23-OCT-2000; 2000US-242612P.		
PR	24-OCT-2000; 2000US-242880P.		
PR	24-OCT-2000; 2000US-242881P.		
PR	29-DEC-2000; 2000US-259028P.		
PR	20-FEB-2001; 2001US-269813P.		
PR	25-MAY-2001; 2001US-286324P.		
PR	29-APR-2001; 2001US-294108P.		
PR	09-JUL-2001; 2001US-303698P.		
PR	16-OCT-2001; 2001US-0981151.		
XX	(CURA-) CURAGEN CORP.		
XX	Edinger S, Gerlach V, MacDougall JK, Malyankar UM, Smithson G;		
PI	Millet J, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shimmets RA;		
PI	Padigaru M, Guo X, Patturajan M, Taupier RJ, Burgess CB;		
PI	Zerhusen BD, Kekuda R, Spytek KA, Gangoli EA, Fernandes ER;		
PI	Gorman L;		
XX	WI: 2002-590434/63.		
DR	N-PSDB; ABS59325.		
XX			
XX	Cytoplasmic, nuclear, membrane bound and secreted polypeptides and		
PT	nucleic acids encoding the polypeptides for diagnosing and treating		
PT	e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and		
PT	diabetes		

XX	PS	Query Match	64.3%; Score 3365; DB 23; Length 856;
XX	CC	Best Local Similarity	71.9%; Pred. No. 1.1e-268;
XX	CC	Matches 660; Conservative 25; Mismatches 59; Indels 174; Gaps 16;	
XX	CC	Sequence 856 AA;	
QY	129	DLVSAYEVDHGRGDYVSHEIMHQRRAVAVSEVSHLRLKGRPHDHLRTSSSLVA	188
DB	19	DLVSAYEVDHGRGDYVSHEIMHQRRAVAVSEVSHLRLKGRPHDHLRTSSSLVA	75
QY	189	---PGFI-----VQTGKTGTSVQTLPPED-----FCFYQ-GSIRS	221
DB	76	NSQFGFLNSVRSWHREHASKRIITNAMLGBSALASTKSCNCFVFFSFGSGMRT	135
QY	222	H-----RNSFSH-----GGKFCBGS-----TRTLKLSQKCPDSVDFRA	257
DB	136	BEADYFLPLPSHLSWKLGRAAQSSPSHVLYKREVLVTSRTWLAHQ---PLHSSDLRL	192
QY	258	AQCAEHNRRFRGRHYKWKPYQVEADLCKLYCAEGDFPFSLNKVKDGTGTCSEDSRN	317
DB	193	GL---PQKQFCGRKK-----	206
QY	318	VCIDGICBLVVSTSAHMQPPKBDLFLPDYKSCSLRHKSLRSHRNEBELVETLVVV	377
DB	207	-----YMPQPKEDLFLPDYKSCSLRHKSLRSHRNEBELVETLVVV	250
QY	378	DKKMQRNHNENITTYVLTILNMYSAFLFKO-----	407
DB	251	DKKMQRNHNENITTYVLTILNMYSAFLFKO-----	310
QY	408	-----GLMKDGTGRDHAILLTGLDICSWNKNEPCDTLGFAPISGMCKYSRST	455
DB	311	DHTLSSFCQWQSGLMGKDGTRDHAILLTGLDICSWNKNEPCDTLGFAPISGMCKYSRST	370
QY	456	INEDTGLGLAFTIAHESGHNFGMTHDGEWNCKSEGNMTSPTLAGNVSFWSPCSROY	515
DB	371	INEDTGLGLAFTIAHESGHNFGMTHDGEWNCKSEGNMTSPTLAGNVSFWSPCSROY	430
QY	516	LHKFLSTQAICLADQPKPKVKEYKPKLPCGLYKANTQCKWQGEKAKLCLMDFKDKIC	575
DB	431	LHKFLSTQAICLADQPKPKVKEYKPKLPCGLYKANTQCKWQGEKAKLCLMDFKDKIC	490
QY	576	KALWCHRIGRCKETKFMPPAASGTICGHDMMCRGQCCKYKIDGPKPTHGHSWSSWSPC	635
DB	491	KALWCHRIGRCKETKFMPPAASGTICGHDMMCRGQCCKYKIDGPKPTHGHSWSSWSPC	549

QY 636 SRTGGVSHRSR-LCTNPKPSHGKFCGSGSTTLKLNCSQKPRDSVDFRAAQAENHS 694
 Db 550 SRTGGVSHRSRQNTSRPSHGKFCGSGSTTLKLNCSQKPRDSVDFRAAQAENHS 609
 QY 695 RFRGRHYKWKPYQVEDQDLCKLYCIAEGFDFPFSSLNKVKDGTGPCSDESRNVCIDGIC 754
 Db 610 RFRGRHYKWKPYQVEDQDLCKLYCIAEGFDFPFSSLNKVKDGTGPCSDESRNVCIDGIC 664
 QY 755 ERVGDNLVGSDAVEDVCGVGNNSACTIHRGLYTKHHHTNOYHMTIPSGARSIRY 814
 Db 665 ERVGDNLVGSDAVEDVCGVGNNSACTIHRGLYTKHHHTNOYHMTIPSGARSIRY 718
 QY 815 ENNVSTSYISRNALRYLNGHWTVDWPGRYKFGCTTDFRYSYNEPENLITATGPTNET 874
 Db 719 ENNVSTSYISRNALRYLNGHWTVDWPGRYKFGCTTDFRYSYNEPENLITATGPTNET 778
 QY 875 LIVELLFQGRNPGVAMEYSPRLGTGTEKQPPAQSFTWAIVRSECVSCGGRCPLPVLLLE 934
 Db 779 LIVELLFQGRNPGVAMEYSPRLGTGTEKQPPAQSFTWAIVRSECVSCGGRCPLPVLLLE 838
 QY 935 AACQPSATAYIALAFLES 952
 Db 839 AACQPSATAYIALAFLES 856

RESULT 11
 AAU98886
 ID AAU98886 standard; Protein; 1236 AA.
 AC AAU98886;
 DT 26-AUG-2002 (first entry)
 DE Human protease PRTS4.
 KW Protease; human; gastrointestinal disorder; gastritis; atherosclerosis;
 KW ulcerative colitis; Reye's syndrome; cardiovascular disorder; enzyme;
 KW hypertension; myocardial infarction; autoimmune disease; AIDS; PRTS;
 KW inflammatory disorder; acquired immunodeficiency syndrome; asthma;
 KW Grave's disease; cell proliferative disorder; hepatitis; psoriasis;
 KW leukaemia; developmental disorder; Cushing's syndrome; impotence;
 KW epithelial disorder; dermatitis; scabies; eczema; neurological disorder;
 KW Parkinson's disease; dementia; Alzheimer's disease; infertility;
 KW Huntington's disease; multiple sclerosis; reproductive disorder.
 OS Homo sapiens.
 XX WO200238744-A2.
 XX 16-MAY-2002.
 XX 18-OCT-2001; 2001WO-US51034.
 XX 18-OCT-2000; 2000US-241573P.
 XX 25-OCT-2000; 2000US-243643P.
 XX 02-NOV-2000; 2000US-245256P.
 XX 13-NOV-2000; 2000US-248395P.
 XX 16-NOV-2000; 2000US-249826P.
 XX 20-NOV-2000; 2000US-252303P.
 XX 01-DEC-2000; 2000US-250981P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Lee EA, Hafalia AJA, Yue H, Lal PG, Yao MG, Lu Y, Wallia NK,
 PI Warren BA, Lu DM, Baughn MR, Delegeane AM, Burford N;
 PI Borowsky ML, Lee S, Xu Y, Griffin JA, Kallick DA, Gandhi AR;
 PI Arvizu C, Ison CH, Tang YT, Azimzai Y, Elliott VS, Swarnakar A;
 PI Rankumar J, Nguyen DB, Tribouley CM, Lo TP, Au-young J;
 PI Thangavelu K, Kearney L;
 XX WPI; 2002-463471/49.
 DR N-PSDB; ABK86137.
 XX

New human proteases useful for diagnosing, preventing or treating anorexia, myocardial infarction, Addison's disease, hepatitis, Cushing's syndrome, eczema, Parkinson's disease, and impotence -
 Claim 1; Page 135-137; 168pp; English.
 This invention relates to the DNA and protein sequences of novel isolated human proteases (PRTS), the protein sequences of the invention are useful for screening a compound for effectiveness as an agonist or antagonist of its activity, the identified agonist and antagonist are useful for treating a disease or condition associated with decreased or overexpression of functional PRTS in a patient. The PRTS protein is also useful as an immunogen for preparing polyclonal or monoclonal antibodies by hybridoma technology. An antibody that binds the PRTS proteins is useful for detection and purification of the proteins and can be used to diagnose a condition or disease associated with expression of PRTS in a subject or in a biological sample. The sequences of the invention are useful for diagnosis, treatment and prevention of gastrointestinal disorders such as gastritis, ulcerative colitis, Reye's syndrome, etc; cardiovascular such as atherosclerosis, hypertension, myocardial infarction, etc; autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), asthma, Grave's disease, etc; cell proliferative disorders such as hepatitis, psoriasis, leukaemia etc; developmental disorders such as Cushing's syndrome; epithelial disorders such as dermatitis, scabies, eczema, etc; neurological disorders such as Parkinson's disease, dementia, Alzheimer's disease, Huntington's disease, multiple sclerosis, etc; or reproductive disorders such as infertility, impotence. The present sequence represents the PRTS4 protein sequence of the invention.
 Sequence 1236 AA;
 Query Match 62.8%; Score 3286; DB 23; Length 1236;
 Best Local Similarity 93.1%; Pred. No. 6.1e-262;
 Matches 591; Conservative 2; Mismatches 0; Indels 42; Gaps 1;
 QY 334 HMPOPPKEDFILPDEYKSCILRHKSLLSRHNEELNVELVVDKMMQNHGHEHTTY 393
 Db 268 YMPOPPKEDFILPDEYKSCILRHKSLLSRHNEELNVELVVDKMMQNHGHEHTTY 327
 QY 394 VLTILNMVSALFKD-----GLMG 411
 Db 328 VLTILNMVSALFKDGTGGNINIAVGLILLEDEQGLVISHHADTLSSFCQWQSLMG 387
 QY 412 KCGTRHDHAILLTGLDICSWKNEPCDTLGFAPISGMSKYSRCTINEDTGLGTAFTAAHE 471
 Db 388 KCGTRHDHAILLTGLDICSWKNEPCDTLGFAPISGMSKYSRCTINEDTGLGTAFTAAHE 447
 QY 472 SGHNFMIHDGEGNMCKKSEGNIMSPTLAGNMGVFSWSPCSROYLHKFLSTAQAICLADQ 531
 Db 448 SGHNFMIHDGEGNMCKKSEGNIMSPTLAGNMGVFSWSPCSROYLHKFLSTAQAICLADQ 507
 QY 532 PKPVKEYKYPEKLPGLYDANTCKQWGEKAKLCMLDFKDKICKALWCHRIKRCETKF 591
 Db 508 PKPVKEYKYPEKLPGLYDANTCKQWGEKAKLCMLDFKDKICKALWCHRIKRCETKF 567
 QY 592 MPAAEGTICGHDWMCRCGQCVKYGDGPKTHGHWSWSSWSPCSRTCCGGVSHRSLCT 651
 Db 568 MPAAEGTICGHDWMCRCGQCVKYGDGPKTHGHWSWSSWSPCSRTCCGGVSHRSLCT 627
 QY 652 NPXPSHGKFCGSGSTTLKLNCSQKPRDSVDFRAAQAENHSRFRGRHYKWKPYTOVE 711
 Db 628 NPXPSHGKFCGSGSTTLKLNCSQKPRDSVDFRAAQAENHSRFRGRHYKWKPYTOVE 687
 QY 712 DQDLCKLYCIAEGDFDFSSLNKVKDGTGPCSDESRNVCIDGICERVGCNVLGSDAVEDV 771
 Db 688 DQDLCKLYCIAEGDFDFSSLNKVKDGTGPCSDESRNVCIDGICERVGCNVLGSDAVEDV 747
 QY 772 CGVCGNNSACTIHRGLYTKHHHTNOYHMTIPSGARSIRIYEMNVSTSYISVRNALRR 831
 Db 748 CGVCGNNSACTIHRGLYTKHHHTNOYHMTIPSGARSIRIYEMNVSTSYISVRNALRR 807
 QY 832 YLLNGHWTVDWPGRYKFGCTTDFRYSYNEPENLITATGPTNETLIVELLFQGRNPGVAME 891

DB 808 YYLNGHWTVDMFGYKFGSTTFDYRRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWE 867
 QY 892 YSMPLRGTEKOPPAQPSYTWAIVRSECVSCGGGR 926
 DB 868 YSMPLRGTEKOPPAQPSYTWAIVRSECVSCGGGQ 902

RESULT 12

AAU79496

ID AAU79496 standard; Protein; 1021 AA.

XX AAU79496;

XX.

DT 15-JUL-2002 (first entry)

DE Human partial ADAMTS protein #1.

XX Human; ADAMTS; cytostatic; antidiabetic; antirheumatic;

XX antithrombotic; antitumor; chronic rheumatoid arthritis; gene therapy;

XX diastolic omentopathy; chronic rheumatoid arthritis; gene therapy;

XX refractory skin ulcer; gastric ulcer; post-operative healing failure;

XX repressin-type 2N-metalloprotease domain; disintegrin-like domain; TSP1;

XX thrombospondin type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;

XX chromosome 5p15.2-15.3; Cri-du-chat syndrome.

XX Homo sapiens.

OS WC200231163-A1.

XX 18-APR-2002.

XX 11-OCT-2001; 2001WO-3P08913.

XX 11-OCT-2000; 2000JP-0311309.

XX 02-APR-2001; 2001JP-0102905.

XX (KAZU-) KAZUSA DNA RES INST FOUND.

XX (MITS-) MITSUBISHI PHARMA CORP.

XX Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;

XX Kanzaki K;

XX WPI; 2002-372277/40.

XX N-PSDB; ABK49821.

XX Human brain-originated ADAMTS family polypeptide and encoded gene,

XX applicable in diagnosis and screening compounds for drug compositions

XX in treating diseases due to e.g. neovascularisation

XX Claim 1; Page 123-127; 172pp; Japanese.

XX The invention relates to a polypeptide belonging to the ADAMTS family is

XX selected from sequences appearing as AAU79496, AAU79497 and AAU79499,

XX a protein that contains the polypeptide, a protein having not less than

XX 50% homology with the amino acid sequence of the polypeptides or a

XX polypeptide modified from any of the polypeptides but with some amino

XX acids deleted, substituted, added or inserted. Also included are the

XX polynucleotides encoding the polypeptides (or their complementary

XX strands or variants), a recombinant vector containing any of the

XX polynucleotides, a transformant which is transformed with the recombinant

XX vector, producing the polypeptide, protein or peptide by culturing the

XX transformant, an antibody that can recognize the polypeptide, protein or

XX peptide and screening compounds to promote or inhibit activity of the

XX polypeptide or protein, or to promote or inhibit expression of the

XX polynucleotide by using the polypeptide, protein, peptide, the

XX polynucleotide, vector, transformant or antibody, particularly in

XX the presence of a test compound for contact before evaluating the

XX activity by measuring signal changes. The polypeptide and encoded gene

XX are applicable in diagnosis and screening compounds for drug compositions

XX in treating diseases due to neovascularisation, diabetic omentopathy,

XX chronic rheumatoid arthritis, angioma, refractory skin and gastric ulcers

XX and post-operative healing failure, including gene therapy.

CC The gene encoding such polypeptide has conserved repressin-type 2N-
 CC metalloprotease domain, disintegrin-like domain and TSP1 (thrombospondin
 CC type 1) domain. Its encoded protein is characterised by high expression
 CC in ovaries, changes in expression dose depending on the sexual cycle, a
 CC decrease in tumour cell and location of the gene on the 5P-syndrome
 CC deletion site on chromosome 5p15.2-15.3 (associated with
 CC Cri-du-chat syndrome). The present sequence represents a ADAMTS
 CC protein of the invention.

XX SQ Sequence 1021 AA;

Query Match 62.6%; Score 3280; DB 23; Length 1021;

Best Local Similarity 92.9%; Pred. No. 1.5e-261;

Matches 590; Conservative 2; Mismatches 1; Indels 42; Gaps 1;

QY 334 HMPQPKREDLFIIPDEYKSLRHRKSLRSHRNEELNVETLVVDKMMQNHGHENITY 393

DB 52 YMPQPKREDLFIIPDEYKSLRHRKSLRSHRNEELNVETLVVDKMMQNHGHENITY 111

QY 394 VLTILANMVSALPKD-----GLMG 411

DB 112 VLTILANMVSALPKDGTGGINNTAIVGLILLEDEQGLVISHHDLTSLSCFOMQGLMG 171

QY 412 KDGTRHDHAILLGLDTCWKNPEPCDTLGFAPISGMSKYRSCITINEDTGLGLAFTIAHE 471

DB 172 KDGTRHDHAILLGLDTCWKNPEPCDTLGFAPISGMSKYRSCITINEDTGLGLAFTIAHE 231

QY 472 SGHNFMIHDGEGNCKKSEGNIMSTLAGRNGVFSWSPCSROYLHKFSTAQAICLAQ 531

DB 232 SGHNFMIHDGEGNCKKSEGNIMSTLAGRNGVFSWSPCSROYLHKFSTAQAICLAQ 291

QY 532 PKPVKEYKYPEKLPGLYDANTCKQFGEKAKLMDLDFKDKOICKALWCHRIKRCETKF 591

DB 292 PKPVKEYKYPEKLPGLYDANTCKQFGEKAKLMDLDFKDKOICKALWCHRIKRCETKF 351

QY 592 MPAAEGTICGHDMMCRGQCCKYKDGPKPTHGHSNDSWSPCSRTCCGGVSHRSLCT 651

DB 352 MPAAEGTICGHDMMCRGQCCKYKDGPKPTHGHSNDSWSPCSRTCCGGVSHRSLCT 411

QY 652 NPKPSHGKFCGSTRTLKLCNSQKCPDSVDPRAAQCAEHNSRRFRGRHYKWKPTQVE 711

DB 412 NPKPSHGKFCGSTRTLKLCNSQKCPDSVDPRAAQCAEHNSRRFRGRHYKWKPTQVE 471

QY 712 DDQLCKLYCIAEGDFEFLSNKVKDGTPCSEPSRNVICDIGICERVGCDNLGSDADEV 771

DB 472 DDQLCKLYCIAEGDFEFLSNKVKDGTPCSEPSRNVICDIGICERVGCDNLGSDADEV 531

QY 772 CGVCGNNSACTTHRGLYTKHHHTNOYHHMTIPSGARSIRIYEMNVSTSYISVRNALRR 831

DB 532 CGVCGNNSACTTHRGLYTKHHHTNOYHHMTIPSGARSIRIYEMNVSTSYISVRNALRR 591

QY 832 YYLNGHWTVDMFGYKFGSTTFDYRRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWE 891

DB 592 YYLNGHWTVDMFGYKFGSTTFDYRRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWE 651

QY 892 YSMPLRGTEKOPPAQPSYTWAIVRSECVSCGGGR 926

DB 652 YSMPLRGTEKOPPAQPSYTWAIVRSECVSCGGGQ 686

RESULT 13

ABG31503

ID ABG31503 standard; Protein; 556 AA.

XX ABG31503;

AC ABG31503;

XX 21-OCT-2002 (first entry)

XX Human ADAM-TS-like Zinc metalloprotease.

XX ADAM TS-like Zinc metalloprotease; TNF-antagonist; psychosis;

XX CD8-antagonist; COPD; congestive heart failure; myocardial infarction;

XX ischaemic heart disease; arrhythmia; hypertension; vascular disease;

KW cerebrovascular disorder; parkinson's disease; motor neuron disease;
 KW dementia; multiple sclerosis; stroke; Alzheimer's disease; dementia;
 KW Huntington's disease; Creutzfeld-Jakob disease; dementia; schizophrenia;
 KW attention deficit disorder.

XX Homo sapiens.
 XX WO200257461-A2.
 XX 25-JUL-2002.
 XX 09-JAN-2002; 2002WO-EP00129.
 XX 18-JAN-2001; 2001US-262034P.
 XX (FARB) BAYER AG.

XX Zhu Z;
 DR WPI; 2002-575457/61.
 DR N-PSDB; ABK90322.
 XX

PT New isolated polynucleotide encoding an ADAM-TS-like zinc
 PT metalloprotease polypeptide, useful for preventing, ameliorating or
 PT treating dysfunctions such as COPD, CNS and cardiovascular disorders -
 XX
 XX Claim 1; Fig 2; 119pp; English.

XX This invention relates to a novel isolated polynucleotide encoding an
 CC ADAM-TS-like zinc metalloprotease polypeptide. The protein of the
 CC invention may have cardiac, asotropic, antiarrhythmic, hypotensive,
 CC antiparkinsonian, nootropic, neuroprotective, cerebroprotective,
 CC anticonvulsant, anti-HIV, neuroleptic and tranquiliser activity.
 CC The DNA sequence of the invention may be used in gene therapy as a
 CC TNF-antagonist or as a CD8-antagonist. The invention also comprises
 CC an expression vector containing the ADAM-TS-like zinc metalloprotease
 CC and reagents that modulate and/or regulate the activity of an ADAM-TS-
 CC like zinc metalloprotease. These reagents can be used in the manufacture
 CC of a medicament for disorders such as COPD, cardiovascular disease
 CC (congestive heart failure, myocardial infarction, ischaemic heart
 CC disease, atrial and ventricular arrhythmias, hypertension and peripheral
 CC vascular diseases) and cerebrovascular disorders (Parkinson's disease,
 CC corticobasal degeneration, motor neuron disease, dementia, multiple
 CC sclerosis, stroke, Alzheimer's disease, Huntington's disease,
 CC Creutzfeld-Jakob dementia, schizophrenia, psychosis and attention
 CC deficit disorders). Pharmaceutical compositions of the invention can
 CC also be used for screening and diagnostic assays. The present
 CC sequence represents the human ADAM-TS-like zinc metalloprotease of the
 CC invention.

XX Sequence 556 AA;
 Query Match 53.7%; Score 2812.5; DB 23; Length 556;
 Best Local Similarity 91.4%; Pred. No. 2.5e-223;
 Matches 508; Conservative 1; Mismatches 2; Indels 45; Gaps 2;

QY 366 NEELNVTLLVVDKMMQNGHENITTYVLTILNWSALFKD----- 407
 DB 1 NEELNVTLLVVDKMMQNGHENITTYVLTILNWSALFKDGTIGGINIAIVGLILLE 60
 QY 408 -----GLMKGCTRHDAHLITGLDICSWKNEPCDTLGFAP 443
 DB 61 DEQPLGVISHHADTLTSFCQWGLMGKGTGRHDAHLITGLDICSWKNEPCDTLGFAP 120
 QY 444 ISGMSKYRSTCTINEDTGLGLAFTIAHESGNHFGMHGDMCKKSEGNIMPTLAGRN 503
 DB 121 ISGMSKYRSTCTINEDTGLGLAFTIAHESGNHFGMHGDMCKKSEGNIMPTLAGRN 180
 QY 504 GVFSWSPCSQRYLHKFLSTAQAICLADQPKPVKEYKYPEKLPGLYDANTQCKWQGEKA 563
 DB 181 GVFSWSPCSQRYLHKFLSTAQAICLADQPKPVKEYKYPEKLPGLYDANTQCKWQGEKA 240
 QY 564 KLCMLDFKKICKALWCHRIGRKCKTFMPAAEGTICGHDWMCRCGGQCVKYGDEGPKPTH 623

DB 241 KLCMLDFKK---ATLWCHRIGRKCKTFMPAAEGTICGHDWMCRCGGQCVKYGDEGPKPTH 297
 QY 624 GHWSWSPCSQRYLHKFLSTAQAICLADQPKPVKEYKYPEKLPGLYDANTQCKWQGEKA 683
 DB 298 GHWSWSPCSQRYLHKFLSTAQAICLADQPKPVKEYKYPEKLPGLYDANTQCKWQGEKA 357
 QY 684 FRAAQCAEHSNRFRGRHYKWKPYTOVEDODLCKLYCIAEGDFFSLSNKVKDGTGPCSE 743
 DB 358 FRAAQCAEHSNRFRGRHYKWKPYTOVEDODLCKLYCIAEGDFFSLSNKVKDGTGPCSE 417
 QY 744 DSRNVCIDGICERVGCDNLVGSDAVEDVCGVCGNNSACTIHRGLYTKHHHTNQYVHMVT 803
 DB 418 DSRNVCIDGICERVGCDNLVGSDAVEDVCGVCGNNSACTIHRGLYTKHHHTNQYVHMVT 477
 QY 804 IPGARSIRIYEMNVSTSYISVRNALRRYVYLNHGMTVDMFGRYKFGSTTDFYRRSYNEPE 863
 DB 478 IPGARSIRIYEMNVSTSYISVRNALRRYVYLNHGMTVDMFGRYKFGSTTDFYRRSYNEPE 537
 QY 864 NLIATGPTNETHLVEL 879
 DB 538 NLIATGPTNETHLVEV 553
 RESULT 14
 ABB81460
 ID ABB81460 standard; Protein; 1221 AA.
 XX
 AC ABB81460;
 XX
 DT 29-AUG-2002 (first entry)
 XX
 DE Human aggrecanase MDS8 protein SEQ ID NO:2.
 XX
 KW Human; aggrecanase; MDS8; enzyme; osteoarthritis; joint disease;
 KW metalloprotease and disintegrin with thrombospondin type 1 repeats 8;
 KW rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic;
 KW chromosome 16.
 XX
 OS Homo sapiens.
 XX
 PN WO200250258-A1.
 XX
 PD 27-JUN-2002.
 XX
 PF 17-DEC-2001; 2001WO-JP11033.
 XX
 PR 18-DEC-2000; 2000JP-0384300.
 XX
 PA (YAMA) YAMANOUCHI PHARM CO LTD.
 XX
 PI Yamaji N, Nishimura K, Abe K;
 XX
 DR WPI; 2002-508794/54.
 DR N-PSDB; ABN89277.
 XX
 PT Aggrecanase and encoded polynucleotide involved in causing joint
 PT diseases, applicable in screening inhibitors as preventives or remedies
 PT for treating the diseases e.g. osteoarthritis and rheumatoid arthritis
 XX
 Claim 1; Page 46-49; 58pp; Japanese.
 XX
 CC The present sequence represents a human protein having aggrecanase
 CC activity, designated MDS8 (metalloprotease and disintegrin with
 CC thrombospondin type 1 repeats 8). The human MDS8 protein has
 CC osteopathic, antirheumatic and antiarthritic activities. The human
 CC MDS8 protein and polynucleotide are applicable in screening
 CC inhibitors as preventives or remedies for treating joint diseases
 CC e.g. osteoarthritis and rheumatoid arthritis. Human MDS8 has been
 CC located to chromosome 16, and more specifically to chromosome
 CC 16q22.3-23.1.
 XX

SQ Sequence 1221 AA;
 Query Match 46.5%; Score 2435.5; DB 23; Length 1221;
 Best Local Similarity 51.3%; Pred. No. 1.1e-191;
 Matches 466; Conservative 118; Mismatches 195; Indels 129; Gaps 17;

Qy 85 ASSSTRGAAGLDGKGRDMEAGNHRSSQQTNTGTENOTLHLVLTQYDLVSAYEVDHGRGYVS 144
 |||:|||||:
 Db 47 ASDSSGASGLN-----DDYVFPVEVDSAGSYIS 77

Qy 145 HEIWHQRRRAVAVSEVSEHLKGRPHFDLRTSSLSVAPGIVQTLGKTGKSV 204
 |||:|||||:
 Db 78 HDILHNGRKKAS-AQNARSSHLFSAFGQSLHLELK-PSAILSSHIVQVLQDGASET 135

Qy 205 QTLPEPDCFYQGSLSHRNPSHGGKFCGEGSTRTLKLCNSQ----KCPRDSVPFRAQC 260
 |||:|||||:
 Db 136 QK-EVQOCFQGFIRND-SSSSVAVSTCAGLSGLIRTRKNEFLISPLPQ-----LLA 186

Qy 261 AEHNSRRFRGRHYKWKPYQVEADLCKLYCIAEGDFPFSLSNKVKDG-----TPCSEDS 315
 |||:|||||:
 Db 187 QEHNYSSPAG-HPHVLVKRTAEKIQRYGYPG-----SGRNYPGYSPSHIPHASQS 238

Qy 316 R-----NVCIDGICELSVVSTSAHMPQPKEDLFLPDEYKSLURHKGSLRS 363
 |||:|||||:
 Db 239 RETEYHRRRLQKHFC-----GRKKYAPKPTEDTYLRFDEYSGSRPRRSAGKS 289

Qy 364 HRNEELNVETLVVDKMMQNHGHEITTYVLTILNMVSALFKDG----- 408
 |||:|||||:
 Db 290 QKG--LVNVELLVADKKNVEXHKGNNVTYILTVMNVSGLFKDGITGSDINVVVSLIL 347

Qy 409 -----LMGDKGTRHDHAILLTGLTDCSWNEPCDITLGF 441
 |||:|||||:
 Db 348 LEQBPGLLINHADQSLNSFCQMSALIGNKGKRDHAILLTGDFICSWNKNEPCDITLGF 407

Qy 442 APISGMCKYKSCITNEDTGLAFTIAHESGHNFMHGDGEGNWKCKSEGNINSPITAG 501
 |||:|||||:
 Db 408 APISGMCKYKSCITNEDTGLAFTIAHESGHNFMHGDGEGNWKCKSEGNINSPITAG 467

Qy 502 RGVFMSWSPCSQRYLHKLFLSTAQAICLAQPKPKVKEYKPEKLGELYDANTOCKWOFGE 561
 |||:|||||:
 Db 468 NNGVFSNSSCSQRYLHKLFLSTPQAGCLVDEPKQAGQKYVDPKLPQIYDADTCKWQFGA 527

Qy 562 KAKLCMLDFKDIKALWCHRIKRCETKFPAAEGTICGHDWCRGQCCVKYGDGPKP 621
 |||:|||||:
 Db 528 KAKLCSLGFKVDICKSLWCHRVHRCETKFPAAEGTICGHDWCRGQCCVKYGDGPKP 587

Qy 622 THGWSWSSWSPCSRTCGGVSHRSRLCTNPKPSHGKFCGEGSTRTLKLCNSQKCRDS 681
 |||:|||||:
 Db 588 IHGWSWSSWSPCSRTCGGVSHRSRLCTNPKPSHGKFCGEGSTRTLKLCNSQKCRDS 647

Qy 682 VDFRAAQCAEHSRRFRGRHYKWKPYQVEADLCKLYCIAEGDFPFSLSNKVKDGTPC 741
 |||:|||||:
 Db 648 LDFRAAQCAEHSRRFRGRHYKWKPYQVEADLCKLYCIAEGDFPFSLSNKVKDGTPC 707

Qy 742 SEDSRNYCIDGICERVCNDVLGSDAIVEDVGVCGNNSACTIHRGLYTKHHTNYHYM 801
 |||:|||||:
 Db 708 SPNKNVDICDGVCLVGDHGLGSAVSDAGVCGKGNSTCKFYKGLYLNQHKANEYYPV 767

Qy 802 VTIPSGARSIRIYEMNVSTSVISVRNALRRYLNGHWTVDWPKYKFSGTFFDYRSVNE 861
 |||:|||||:
 Db 768 VILPAGARSIBELQWSSSVLAVRSLSQKYLITGWSIDWPGFPFAGTTFEQRGNFR 827

Qy 862 PENLIATGFTNETLIVELLFGNRPNGVAWEYSMPRL--GTEKQPPA--QPSYTWAIYRSE 917
 |||:|||||:
 Db 828 PERLYAPGPTNETLVFELLMQGNPGIAWKALPKVWNGT---PPATKRPAYTWSIQSE 884

Qy 918 CSVSCGGG 925
 |||:|||||:
 Db 885 CSVSCGGG 892

RESULT 15
 ABJ19379

ID XX ABJ19379 standard; Protein; 1162 AA.
 AC ABJ19379;
 XX 28-MAR-2003 (first entry)
 DT XX NOXV related protein SEQ ID No 124.
 DE XX
 DE XX
 KW Antidiabetic; anorectic; virucide; antibacterial; fungicide; nootropic;
 KW protozoa; neuroprotective; antiparkinsonian; antilipase; antilipase;
 KW NOXV-associated disorder; metabolic disorder; diabetes; anorexia;
 KW obesity; infectious disease; cancer-associated cachexia; immune disorder;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW haematopoietic disorder; cancer; dyslipidaemia; metabolic disturbance;
 KW neurogenesis; cell differentiation; cell proliferation; haematopoiesis;
 KW wound healing; angiogenesis; gene therapy; chromosome mapping;
 KW tissue typing; preventive medicine; pharmacogenomic; NOXV.
 XX
 OS Unidentified.
 OS
 XX WC300299062-A2.
 XX
 PD 12-DEC-2002.
 XX
 XX
 XX 04-JUN-2002; 2002WC-US17559.
 XX
 PR 04-JUN-2001; 2001US-295607P.
 PR 06-JUN-2001; 2001US-296404P.
 PR 06-JUN-2001; 2001US-296418P.
 PR 07-JUN-2001; 2001US-296575P.
 PR 11-JUN-2001; 2001US-297414P.
 PR 12-JUN-2001; 2001US-297567P.
 PR 12-JUN-2001; 2001US-297573P.
 PR 14-JUN-2001; 2001US-298285P.
 PR 15-JUN-2001; 2001US-298528P.
 PR 15-JUN-2001; 2001US-298556P.
 PR 21-JUN-2001; 2001US-299949P.
 PR 22-JUN-2001; 2001US-300177P.
 PR 28-JUN-2001; 2001US-301530P.
 PR 28-JUN-2001; 2001US-301550P.
 PR 03-JUL-2001; 2001US-302951P.
 PR 12-SEP-2001; 2001US-318771P.
 PR 25-SEP-2001; 2001US-324687P.
 PR 24-OCT-2001; 2001US-339266P.
 PR 16-NOV-2001; 2001US-337524P.
 PR 14-DEC-2001; 2001US-341143P.
 PR 21-FEB-2002; 2002US-358643P.
 PR 21-FEB-2002; 2002US-359151P.
 PR 28-FEB-2002; 2002US-361195P.
 PR 05-MAR-2002; 2002US-361964P.
 PR 10-APR-2002; 2002US-371346P.
 PR 10-APR-2002; 2002US-371523P.
 PR 03-JUN-2002; 2002US-0161493.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Anderson DW, Zerhusen BD, Li L, Zhong M, Casman SJ, Gerlach VL;
 PI Shimkets RA, Gorman L, Pena CEA, Kekuda R, Patturajan M;
 PI Spyttek KA, Leite MW, Rastelli L, MacDougall JR, Taupier RJ, Guo X;
 PI Miller CE, Shenoy SG, Rjalt T, Voss EZ, Boldog FT, Malyankar UM;
 PI Padigaru M, Ji W, Smithson G, Edinger SR, Millet i, Ellerman K;
 XX WPI; 2003-140607/13.
 DR N-PSDB; ABT16078.
 XX
 XX New isolated NOXV polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOXV-associated disorders, e.g.
 PT obesity, cancer, Parkinson's disease, infections, immune disorders, or
 PT various dyslipidemias
 XX
 PS Example 39; Page 233-234; 461pp; English.
 XX
 XX The invention relates to an isolated polypeptide comprising any of the 36

86-1370 residue amino acid sequences, given in the specification, a mature form of them, or a sequence that is at least 95 % identical to, or having one or more conservative amino acid substitutions in one of the 36 amino acid sequences. The polypeptides, nucleic acid molecules and antibodies of the invention are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOX-associated disorder. The nucleic acid molecules, polypeptides and antibodies are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer-associated cachexia, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, cancer and various dyslipidaemias, or metabolic disturbances associated with obesity, metabolic x syndrome, and wasting disorders. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, haematopoiesis, wound healing and angiogenesis, in gene therapy, in generation of antibodies that bind immunospecifically to NOX substances for use in therapeutic or diagnostic methods. The nucleic acids are further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. This sequence represents a NOX related protein of the invention.

xx Sequence 1162 AA;

Query Match 45.1%; Score 2364; DB 24; Length 1162;
Best Local Similarity 50.3%; Pred. No. 8.2e-186;
Matches 460; Conservative 116; Mismatches 197; Indels 142; Gaps 20;

Qy 85 ASSSTRGAAGLDGKGRDMEAGNHRSSQQTNTGTENQTLHLVLTQYDLVSAYEVDHRGDYVS 144
Db 47 ASDSSGASGLN-----DDYVFTPEVDGAGSYIS 77

Qy 145 HEIMHQRRRAVAVSEVESHLLKGRPHFDLRTSSSLVAPGFIVQLGKTGTSKV 204
Db 78 HDILHNGKKKS-AQNARSSLHYRPSAFQGBLHLELK-PSAILSSHFIQVLGKDGAS 135

Qy 205 QTLPPEDFCFYQGLSRHNRGSPSHGGKFCGEGSTRTLKCNQ-----KCPRDSVDFRAAQ 260
Db 136 QK-PEVQCQCFYQGFIRND-SSSSVAVSTCAGLSGLIRKNEFLISLPQ-----LLA 186

Qy 261 AEHNSRRFRGRHYKWKPYTQVEADLCKLYCIAEGDFDFPSLSNKKVDG-----TPCSEDS 315
Db 187 QEHNYSPAG-HPHVLYKRTAEKIQRYGYPG-----SGRNYGYPSPHIFASQS 238

Qy 316 R-----NVCIDIGELSVVSTSAHPPQPPKEDLILPDEYKSLRHKRSLRS 363
Db 239 RETEYHRRLOKHFC-----GRRKKYAPKPTEDTYLRFDEYGGSGRPRRSAGKS 289

Qy 364 HRNEELNVETLVVDKQKMONHGHENITTYVLTILNMVSALFKDG----- 408
Db 290 QKG--LNVETLVVDKKNVEXHKGKGNVTYILTVMNVYGLFKDGTIGSDINVVVSLIL 347

Qy 409 -----LMGKDGTRHDHAILLTGLDICSWNKNEPCDTLGF 441
Db 348 LEQEPGGLINHHADQSLNSFCQMSALIGNKGKRDHAILLTGTFDICSWNKNEPCDTLGF 407

Qy 442 APIQCMCKYRSCTINEDTGLAFTIAHESGHNFGMTHDGEWNCKSEGNIMSPITLAG 501
Db 408 APIQCMCKYRSCTINEDTGLAFTIAHESGHNFGMTHDGEWNCKSEGNIMSPITLAG 467

Qy 502 RNVFWSWPCSRQYLHKFLSTAQAICLADQPKPKVKEYKYPKELFGLYDANTQCKWQFGE 561
Db 468 NNGVFSWSWPCSRQYLKFLSTPQAGCLVDEPKAQKYPKDKLPQIYDADTQCKWQFGE 527

Qy 562 KAKLCHLPFKDICIHALWCHIRGKRCETKFPAAEGTICGHDMCRGGQCVKYGDEGPKP 621
Db 528 KAKLSLGLFVKDICKLWCHRVGHRCEKTFKPAAGTVCGLSMWCRQGCQCVKFGELGPRP 587

Qy 622 THGWSWSSWPCSRCTCGGVSHRSRLCTNPKPSHGKFCGEGSTRTLKCNLSQKCPRDS 681
Db 588 IHGWSAWSKWSERCTCGGVVKFQERHCNNPKQYGGLCFPGSSRIYQLCNINPCNENS 647

Qy 682 VDFRAQCAEHSRRFRGRHYKWKPYTQVEADLCKLYCIAEGDFDFPSLSNKKVDGTPC 741
Db 648 LDFRAQCAEAYSNSKPFGRWFYQWKPYTKVEEDRCKLYCKAENFEFFFAWSGKVDGTPC 707

Qy 742 SEDSRNVCIDGICERVGCNDVLGSDADEVCGVCGNNSACTIHRGLYTKHHHTNQYYHM 801
Db 708 SPKNVDVICDGVCELVCCHDELGSKAVSACGVCKGDNSTCKFYKGLYLQHKANEYYPV 767

Qy 802 VTIPSGARSIRIYEMNVSTSYISVRNALRRYYLNGHWTVDWPGRYKFSGTTFDYRKSNE 861
Db 768 VLIIPAGARSIEIQELQVSSSYLAVRSLSKQYLYLTGWSIDWPGCFPFAGTTFEYQKSFNR 827

Qy 862 PENLIATGPTNETLIVELLFOGRRPGVAEYSMPRL--GTEKOPPA--OPSYT-----W 911
Db 828 PERLYAPGPTNETLI---LMQGNKPGIAWKYALPKYMNGT---PPATKRPAITCWNMPGEW 881

Qy 912 AIVRSECSVSCGGGR 926
Db 882 ----STCSKACAGGQ 892

Search completed: October 28, 2003, 23:40:52
Job time : 86 secs


```

Db 361 TVCDPSRSCSVTDDGLOAFTTAHELGHVFNPHD--DAKHCASLNGVTGSHLWASMLS 419
Qy 501 GRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPKVKEYKPEKLPGLDYDANTCKWQFG 560
Db 420 SLDSQSPWSPCSAYVWVTFSLDNGHGECLMD--KPQNPICKLPSDLPGTLIDANRQCOPTFG 477
Qy 561 EKAKLMLDFKXDKICALWCHRIQK-----CETKFPAAEGTICGHDMWCHRGOCVKYG 615
Db 478 BESKHC--PDAASTCTTLMC--TGTSGLLVQCTKHPWADGTSCGEGKWCVSGKCVNT 533
Qy 616 D--EGPKPTHGWSWSSWSPCSRCTCGGVSHRSRLCTNPKPSHGKFCBGSTRTLKLCN 673
Db 534 DMKHATPVGWGWGPGWAGCDSRTCCTCGGVQVQVTRMCDNPVPGKVKYCEGKVRYSRSCN 593
Qy 674 SQKCP--RDSVDFAAQAEBHN--SRRFRGRH--YKWKPK--YTQVEDQDCKLYCIAEGFDF 727
Db 594 IEDCPDNNGKTFREQCEAHNEFSKASFGNEPTVETWTPKYAGVSPKDRCKLTCEAKGIGY 653
Qy 728 FFSLSNKVKDGTCTCEDSRNVCIDGICERVGCDNVLGSDAEDVCGVCGNNSACTIHRG 787
Db 654 FVLQPKVVDGTFPCSPDSTSVQVQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKRMGS 713
Qy 788 LVTKHHTNQYVH--MVTIPSGARSIRIYEMNV-----STSYISVENALRRYVINGHWTV 841
Db 714 IVTS--TPRGYHDIVTFAGATNIEVKHNRQGRNNGSPLAIRAADGTYLINGNFTLS 770
Qy 842 WPGR--YKFSGTTFDYRRSYNEPENLIATOPTNETLIVELLFOGR--NPGVAWEYSMPRLG 898
Db 771 TLEQDLTYKGTVLYRSGSSAALERIRFSPLKEPTIQLVMVGHALRPKIKETPTFMKK-- 828
Qy 899 TEKQPPAQPST--WAIVR--SECVSOGGG 925
Db 829 KTESFNAITPSEWVIEENGELSKTCSG 857

RESULT 2
T21371
hypothetical protein F25H8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
A:Accession: T21371, T24896
R:GajadstV, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19413
A:Accession: T21371
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone F25H8
R:GajadstV, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19949
A:Accession: T24896
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <W12>
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone T13H10
C:Genetics:
A:Gene: CESP:F25H8.3
A:Map position: 4
A:Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 811/1; 812/1; 813/1; 814/1; 815/1; 816/1; 817/1; 818/1; 819/1; 820/1; 821/1; 822/1; 823/1; 824/1; 825/1; 826/1; 827/1; 828/1; 829/1; 830/1; 831/1; 832/1; 833/1; 834/1; 835/1; 836/1; 837/1; 838/1; 839/1; 840/1; 841/1; 842/1; 843/1; 844/1; 845/1; 846/1; 847/1; 848/1; 849/1; 850/1; 851/1; 852/1; 853/1; 854/1; 855/1; 856/1; 857/1; 858/1; 859/1; 860/1; 861/1; 862/1; 863/1; 864/1; 865/1; 866/1; 867/1; 868/1; 869/1; 870/1; 871/1; 872/1; 873/1; 874/1; 875/1; 876/1; 877/1; 878/1; 879/1; 880/1; 881/1; 882/1; 883/1; 884/1; 885/1; 886/1; 887/1; 888/1; 889/1; 890/1; 891/1; 892/1; 893/1; 894/1; 895/1; 896/1; 897/1; 898/1; 899/1; 900/1; 901/1; 902/1; 903/1; 904/1; 905/1; 906/1; 907/1; 908/1; 909/1; 910/1; 911/1; 912/1; 913/1; 914/1; 915/1; 916/1; 917/1; 918/1; 919/1; 920/1; 921/1; 922/1; 923/1; 924/1; 925/1; 926/1; 927/1; 928/1; 929/1; 930/1; 931/1; 932/1; 933/1; 934/1; 935/1; 936/1; 937/1; 938/1; 939/1; 940/1; 941/1; 942/1; 943/1; 944/1; 945/1; 946/1; 947/1; 948/1; 949/1; 950/1; 951/1; 952/1; 953/1; 954/1; 955/1; 956/1; 957/1; 958/1; 959/1; 960/1; 961/1; 962/1; 963/1; 964/1; 965/1; 966/1; 967/1; 968/1; 969/1; 970/1; 971/1; 972/1; 973/1; 974/1; 975/1; 976/1; 977/1; 978/1; 979/1; 980/1; 981/1; 982/1; 983/1; 984/1; 985/1; 986/1; 987/1; 988/1; 989/1; 990/1; 991/1; 992/1; 993/1; 994/1; 995/1; 996/1; 997/1; 998/1; 999/1; 1000/1; 1001/1; 1002/1; 1003/1; 1004/1; 1005/1; 1006/1; 1007/1; 1008/1; 1009/1; 1010/1; 1011/1; 1012/1; 1013/1; 1014/1; 1015/1; 1016/1; 1017/1; 1018/1; 1019/1; 1020/1; 1021/1; 1022/1; 1023/1; 1024/1; 1025/1; 1026/1; 1027/1; 1028/1; 1029/1; 1030/1; 1031/1; 1032/1; 1033/1; 1034/1; 1035/1; 1036/1; 1037/1; 1038/1; 1039/1; 1040/1; 1041/1; 1042/1; 1043/1; 1044/1; 1045/1; 1046/1; 1047/1; 1048/1; 1049/1; 1050/1; 1051/1; 1052/1; 1053/1; 1054/1; 1055/1; 1056/1; 1057/1; 1058/1; 1059/1; 1060/1; 1061/1; 1062/1; 1063/1; 1064/1; 1065/1; 1066/1; 1067/1; 1068/1; 1069/1; 1070/1; 1071/1; 1072/1; 1073/1; 1074/1; 1075/1; 1076/1; 1077/1; 1078/1; 1079/1; 1080/1; 1081/1; 1082/1; 1083/1; 1084/1; 1085/1; 1086/1; 1087/1; 1088/1; 1089/1; 1090/1; 1091/1; 1092/1; 1093/1; 1094/1; 1095/1; 1096/1; 1097/1; 1098/1; 1099/1; 1100/1; 1101/1; 1102/1; 1103/1; 1104/1; 1105/1; 1106/1; 1107/1; 1108/1; 1109/1; 1110/1; 1111/1; 1112/1; 1113/1; 1114/1; 1115/1; 1116/1; 1117/1; 1118/1; 1119/1; 1120/1; 1121/1; 1122/1; 1123/1; 1124/1; 1125/1; 1126/1; 1127/1; 1128/1; 1129/1; 1130/1; 1131/1; 1132/1; 1133/1; 1134/1; 1135/1; 1136/1; 1137/1; 1138/1; 1139/1; 1140/1; 1141/1; 1142/1; 1143/1; 1144/1; 1145/1; 1146/1; 1147/1; 1148/1; 1149/1; 1150/1; 1151/1; 1152/1; 1153/1; 1154/1; 1155/1; 1156/1; 1157/1; 1158/1; 1159/1; 1160/1; 1161/1; 1162/1; 1163/1; 1164/1; 1165/1; 1166/1; 1167/1; 1168/1; 1169/1; 1170/1; 1171/1; 1172/1; 1173/1; 1174/1; 1175/1; 1176/1; 1177/1; 1178/1; 1179/1; 1180/1; 1181/1; 1182/1; 1183/1; 1184/1; 1185/1; 1186/1; 1187/1; 1188/1; 1189/1; 1190/1; 1191/1; 1192/1; 1193/1; 1194/1; 1195/1; 1196/1; 1197/1; 1198/1; 1199/1; 1200/1; 1201/1; 1202/1; 1203/1; 1204/1; 1205/1; 1206/1; 1207/1; 1208/1; 1209/1; 1210/1; 1211/1; 1212/1; 1213/1; 1214/1; 1215/1; 1216/1; 1217/1; 1218/1; 1219/1; 1220/1; 1221/1; 1222/1; 1223/1; 1224/1; 1225/1; 1226/1; 1227/1; 1228/1; 1229/1; 1230/1; 1231/1; 1232/1; 1233/1; 1234/1; 1235/1; 1236/1; 1237/1; 1238/1; 1239/1; 1240/1; 1241/1; 1242/1; 1243/1; 1244/1; 1245/1; 1246/1; 1247/1; 1248/1; 1249/1; 1250/1; 1251/1; 1252/1; 1253/1; 1254/1; 1255/1; 1256/1; 1257/1; 1258/1; 1259/1; 1260/1; 1261/1; 1262/1; 1263/1; 1264/1; 1265/1; 1266/1; 1267/1; 1268/1; 1269/1; 1270/1; 1271/1; 1272/1; 1273/1; 1274/1; 1275/1; 1276/1; 1277/1; 1278/1; 1279/1; 1280/1; 1281/1; 1282/1; 1283/1; 1284/1; 1285/1; 1286/1; 1287/1; 1288/1; 1289/1; 1290/1; 1291/1; 1292/1; 1293/1; 1294/1; 1295/1; 1296/1; 1297/1; 1298/1; 1299/1; 1300/1; 1301/1; 1302/1; 1303/1; 1304/1; 1305/1; 1306/1; 1307/1; 1308/1; 1309/1; 1310/1; 1311/1; 1312/1; 1313/1; 1314/1; 1315/1; 1316/1; 1317/1; 1318/1; 1319/1; 1320/1; 1321/1; 1322/1; 1323/1; 1324/1; 1325/1; 1326/1; 1327/1; 1328/1; 1329/1; 1330/1; 1331/1; 1332/1; 1333/1; 1334/1; 1335/1; 1336/1; 1337/1; 1338/1; 1339/1; 1340/1; 1341/1; 1342/1; 1343/1; 1344/1; 1345/1; 1346/1; 1347/1; 1348/1; 1349/1; 1350/1; 1351/1; 1352/1; 1353/1; 1354/1; 1355/1; 1356/1; 1357/1; 1358/1; 1359/1; 1360/1; 1361/1; 1362/1; 1363/1; 1364/1; 1365/1; 1366/1; 1367/1; 1368/1; 1369/1; 1370/1; 1371/1; 1372/1; 1373/1; 1374/1; 1375/1; 1376/1; 1377/1; 1378/1; 1379/1; 1380/1; 1381/1; 1382/1; 1383/1; 1384/1; 1385/1; 1386/1; 1387/1; 1388/1; 1389/1; 1390/1; 1391/1; 1392/1; 1393/1; 1394/1; 1395/1; 1396/1; 1397/1; 1398/1; 1399/1; 1400/1; 1401/1; 1402/1; 1403/1; 1404/1; 1405/1; 1406/1; 1407/1; 1408/1; 1409/1; 1410/1; 1411/1; 1412/1; 1413/1; 1414/1; 1415/1; 1416/1; 1417/1; 1418/1; 1419/1; 1420/1; 1421/1; 1422/1; 1423/1; 1424/1; 1425/1; 1426/1; 1427/1; 1428/1; 1429/1; 1430/1; 1431/1; 1432/1; 1433/1; 1434/1; 1435/1; 1436/1; 1437/1; 1438/1; 1439/1; 1440/1; 1441/1; 1442/1; 1443/1; 1444/1; 1445/1; 1446/1; 1447/1; 1448/1; 1449/1; 1450/1; 1451/1; 1452/1; 1453/1; 1454/1; 1455/1; 1456/1; 1457/1; 1458/1; 1459/1; 1460/1; 1461/1; 1462/1; 1463/1; 1464/1; 1465/1; 1466/1; 1467/1; 1468/1; 1469/1; 1470/1; 1471/1; 1472/1; 1473/1; 1474/1; 1475/1; 1476/1; 1477/1; 1478/1; 1479/1; 1480/1; 1481/1; 1482/1; 1483/1; 1484/1; 1485/1; 1486/1; 1487/1; 1488/1; 1489/1; 1490/1; 1491/1; 1492/1; 1493/1; 1494/1; 1495/1; 1496/1; 1497/1; 1498/1; 1499/1; 1500/1; 1501/1; 1502/1; 1503/1; 1504/1; 1505/1; 1506/1; 1507/1; 1508/1; 1509/1; 1510/1; 1511/1; 1512/1; 1513/1; 1514/1; 1515/1; 1516/1; 1517/1; 1518/1; 1519/1; 1520/1; 1521/1; 1522/1; 1523/1; 1524/1; 1525/1; 1526/1; 1527/1; 1528/1; 1529/1; 1530/1; 1531/1; 1532/1; 1533/1; 1534/1; 1535/1; 1536/1; 1537/1; 1538/1; 1539/1; 1540/1; 1541/1; 1542/1; 1543/1; 1544/1; 1545/1; 1546/1; 1547/1; 1548/1; 1549/1; 1550/1; 1551/1; 1552/1; 1553/1; 1554/1; 1555/1; 1556/1; 1557/1; 1558/1; 1559/1; 1560/1; 1561/1; 1562/1; 1563/1; 1564/1; 1565/1; 1566/1; 1567/1; 1568/1; 1569/1; 1570/1; 1571/1; 1572/1; 1573/1; 1574/1; 1575/1; 1576/1; 1577/1; 1578/1; 1579/1; 1580/1; 1581/1; 1582/1; 1583/1; 1584/1; 1585/1; 1586/1; 1587/1; 1588/1; 1589/1; 1590/1; 1591/1; 1592/1; 1593/1; 1594/1; 1595/1; 1596/1; 1597/1; 1598/1; 1599/1; 1600/1; 1601/1; 1602/1; 1603/1; 1604/1; 1605/1; 1606/1; 1607/1; 1608/1; 1609/1; 1610/1; 1611/1; 1612/1; 1613/1; 1614/1; 1615/1; 1616/1; 1617/1; 1618/1; 1619/1; 1620/1; 1621/1; 1622/1; 1623/1; 1624/1; 1625/1; 1626/1; 1627/1; 1628/1; 1629/1; 1630/1; 1631/1; 1632/1; 1633/1; 1634/1; 1635/1; 1636/1; 1637/1; 1638/1; 1639/1; 1640/1; 1641/1; 1642/1; 1643/1; 1644/1; 1645/1; 1646/1; 1647/1; 1648/1; 1649/1; 1650/1; 1651/1; 1652/1; 1653/1; 1654/1; 1655/1; 1656/1; 1657/1; 1658/1; 1659/1; 1660/1; 1661/1; 1662/1; 1663/1; 1664/1; 1665/1; 1666/1; 1667/1; 1668/1; 1669/1; 1670/1; 1671/1; 1672/1; 1673/1; 1674/1; 1675/1; 1676/1; 1677/1; 1678/1; 1679/1; 1680/1; 1681/1; 1682/1; 1683/1; 1684/1; 1685/1; 1686/1; 1687/1; 1688/1; 1689/1; 1690/1; 1691/1; 1692/1; 1693/1; 1694/1; 1695/1; 1696/1; 1697/1; 1698/1; 1699/1; 1700/1; 1701/1; 1702/1; 1703/1; 1704/1; 1705/1; 1706/1; 1707/1; 1708/1; 1709/1; 1710/1; 1711/1; 1712/1; 1713/1; 1714/1; 1715/1; 1716/1; 1717/1; 1718/1; 1719/1; 1720/1; 1721/1; 1722/1; 1723/1; 1724/1; 1725/1; 1726/1; 1727/1; 1728/1; 1729/1; 1730/1; 1731/1; 1732/1; 1733/1; 1734/1; 1735/1; 1736/1; 1737/1; 1738/1; 1739/1; 1740/1; 1741/1; 1742/1; 1743/1; 1744/1; 1745/1; 1746/1; 1747/1; 1748/1; 1749/1; 1750/1; 1751/1; 1752/1; 1753/1; 1754/1; 1755/1; 1756/1; 1757/1; 1758/1; 1759/1; 1760/1; 1761/1; 1762/1; 1763/1; 1764/1; 1765/1; 1766/1; 1767/1; 1768/1; 1769/1; 1770/1; 1771/1; 1772/1; 1773/1; 1774/1; 1775/1; 1776/1; 1777/1; 1778/1; 1779/1; 1780/1; 1781/1; 1782/1; 1783/1; 1784/1; 1785/1; 1786/1; 1787/1; 1788/1; 1789/1; 1790/1; 1791/1; 1792/1; 1793/1; 1794/1; 1795/1; 1796/1; 1797/1; 1798/1; 1799/1; 1800/1; 1801/1; 1802/1; 1803/1; 1804/1; 1805/1; 1806/1; 1807/1; 1808/1; 1809/1; 1810/1; 1811/1; 1812/1; 1813/1; 1814/1; 1815/1; 1816/1; 1817/1; 1818/1; 1819/1; 1820/1; 1821/1; 1822/1; 1823/1; 1824/1; 1825/1; 1826/1; 1827/1; 1828/1; 1829/1; 1830/1; 1831/1; 1832/1; 1833/1; 1834/1; 1835/1; 1836/1; 1837/1; 1838/1; 1839/1; 1840/1; 1841/1; 1842/1; 1843/1; 1844/1; 1845/1; 1846/1; 1847/1; 1848/1; 1849/1; 1850/1; 1851/1; 1852/1; 1853/1; 1854/1; 1855/1; 1856/1; 1857/1; 1858/1; 1859/1; 1860/1; 1861/1; 1862/1; 1863/1; 1864/1; 1865/1; 1866/1; 1867/1; 1868/1; 1869/1; 1870/1; 1871/1; 1872/1; 1873/1; 1874/1; 1875/1; 1876/1; 1877/1; 1878/1; 1879/1; 1880/1; 1881/1; 1882/1; 1883/1; 1884/1; 1885/1; 1886/1; 1887/1; 1888/1; 1889/1; 1890/1; 1891/1; 1892/1; 1893/1; 1894/1; 1895/1; 1896/1; 1897/1; 1898/1; 1899/1; 1900/1; 1901/1; 1902/1; 1903/1; 1904/1; 1905/1; 1906/1; 1907/1; 1908/1; 1909/1; 1910/1; 1911/1; 1912/1; 1913/1; 1914/1; 1915/1; 1916/1; 1917/1; 1918/1; 1919/1; 1920/1; 1921/1; 1922/1; 1923/1; 1924/1; 1925/1; 1926/1; 1927/1; 1928/1; 1929/1; 1930/1; 1931/1; 1932/1; 1933/1; 1934/1; 1935/1; 1936/1; 1937/1; 1938/1; 1939/1; 1940/1; 1941/1; 1942/1; 1943/1; 1944/1; 1945/1; 1946/1; 1947/1; 1948/1; 1949/1; 1950/1; 1951/1; 1952/1; 1953/1; 1954/1; 1955/1; 1956/1; 1957/1; 1958/1; 1959/1; 1960/1; 1961/1; 1962/1; 1963/1; 1964/1; 1965/1; 1966/1; 1967/1; 1968/1; 1969/1; 1970/1; 1971/1; 1972/1; 1973/1; 1974/1; 1975/1; 1976/1; 1977/1; 1978/1; 1979/1; 1980/1; 1981/1; 1982/1; 1983/1; 1984/1; 1985/1; 1986/1; 1987/1; 1988/1; 1989/1; 1990/1; 1991/1; 1992/1; 1993/1; 1994/1; 1995/1; 1996/1; 1997/1; 1998/1; 1999/1; 2000/1; 2001/1; 2002/1; 2003/1; 2004/1; 2005/1; 2006/1; 2007/1; 2008/1; 2009/1; 2010/1; 2011/1; 2012/1; 2013/1; 2014/1; 2015/1; 2016/1; 2017/1; 2018/1; 2019/1; 2020/1; 2021/1; 2022/1; 2023/1; 2024/1; 2025/1; 2026/1; 2027/1; 2028/1; 2029/1; 2030/1; 2031/1; 2032/1; 2033/1; 2034/1; 2035/1; 2036/1; 2037/1; 2038/1; 2039/1; 2040/1; 2041/1; 2042/1; 2043/1; 2044/1; 2045/1; 2046/1; 2047/1; 2048/1; 2049/1; 2050/1; 2051/1; 2052/1; 2053/1; 2054/1; 2055/1; 2056/1; 2057/1; 2058/1; 2059/1; 2060/1; 2061/1; 2062/1; 2063/1; 2064/1; 2065/1; 2066/1; 2067/1; 2068/1; 2069/1; 2070/1; 2071/1; 2072/1; 2073/1; 2074/1; 2075/1; 2076/1; 2077/1; 2078/1; 2079/1; 2080/1; 2081/1; 2082/1; 2083/1; 2084/1; 2085/1; 2086/1; 2087/1; 2088/1; 2089/1; 2090/1; 2091/1; 2092/1; 2093/1; 2094/1; 2095/1; 2096/1; 2097/1; 2098/1; 2099/1; 2100/1; 2101/1; 2102/1; 2103/1; 2104/1; 2105/1; 2106/1; 2107/1; 2108/1; 2109/1; 2110/1; 2111/1; 2112/1; 2113/1; 2114/1; 2115/1; 2116/1; 2117/1; 2118/1; 2119/1; 2120/1; 2121/1; 2122/1; 2123/1; 2124/1; 2125/1; 2126/1; 2127/1; 2128/1; 2129/1; 2130/1; 2131/1; 2132/1; 2133/1; 2134/1; 2135/1; 2136/1; 2137/1; 2138/1; 2139/1; 2140/1; 2141/1; 2142/1; 2143/1; 2144/1; 2145/1; 2146/1; 2147/1; 2148/1; 2149/1; 2150/1; 2151/1; 2152/1; 2153/1; 2154/1; 2155/1; 2156/1; 2157/1; 2158/1; 2159/1; 2160/1; 2161/1; 2162/1; 2163/1; 2164/1; 2165/1; 2166/1; 2167/1; 2168/1; 2169/1; 2170/1; 2171/1; 2172/1; 2173/1; 2174/1; 2175/1; 2176/1; 2177/1; 2178/1; 2179/1; 2180/1; 2181/1; 2182/1; 2183/1; 2184/1; 2185/1; 2186/1; 2187/1; 2188/1; 2189/1; 2190/1; 2191/1; 2192/1; 2193/1; 2194/1; 2195/1; 2196/1; 2197/1; 2198/1; 2199/1; 2200/1; 2201/1; 2202/1; 2203/1; 2204/1; 2205/1; 2206/1; 2207/1; 2208/1; 2
```


A:Molecule type: mRNA
A:Residues: 1-1205 <COL>
A:Cross-references: EMBL:X6389; NID:e990769; PID:e228215; PIDN:CAA6523.1
A:Experimental source: skin
C:Genetics:
A:Gene: PC I-NP
C:Function:
A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to
C:Keywords: hydrolase; metalloproteinase

Query Match 18.1%; Score 948; DB 2; Length 1205;
Best Local Similarity 29.7%; Pred. No. 3e-60;
Matches 260; Conservative 131; Mismatches 333; Indels 152; Gaps 30;
QY 137 DHRGDDVSHSEIMHQR-----RRRAVAVSEVESL-----HLRLKGRPHDFHM 178
DB 57 DAQGLVSHVVSAAATAPAGVTRRAAPAIPLSGGSEEDGGRLFYVIVFG--RDHL 114
QY 179 DLRSSSLVAPGFTVQTLGKTGYSVQTLPPEDCFYQGLSRSHRNPSFHGKFCGSTR 238
DB 115 RLRLNARLAVPAGATVEQGESGATRV--PLLGLTCLYVDVAGLAESSVALSNCDDLGLAG 172
QY 239 TLKLCNOKCPDRSDVDFRAACCAEHNSRRFRGR-HYKW-KPYTQVEADLCKLYCIAEGFD 296
DB 173 LIRNEEEFFIEPLEKGLAAKEA-----QGRVHVHRTTSRPPPL----- 215
QY 297 FFFSLSNKVKDGTPTCSDESRNVICDIGICELSVVSTSAHMFQPKKEDLFIIPDEYKSLRH 356
DB 216 -----GGQALDT-GISADSLSL-----RALGVLEERVNSRRR 250
QY 357 KESLLSRHNEELNVEVLVVDKMMQNHGHEITTVTLILNMVSALFKDGLMG----- 411
DB 251 MR-----RHAADDDYIEVLGVDDSVQVFGTEHVQKYLTLNMVINEIYHDESLGAHNV 307
QY 412 -----KQGT-----RHDHAILLTGLDICS 430
DB 308 VLVRILLYSYKMSLIEIGNPSQSLNVCNVAILOQKPTDDEYDHAIFTRQF-- 365
QY 431 WKNEPCDTLGFAPISGCMKYRSTINEDTGLGLAFTIAHSGHNFQMIHDEGNCM--K 488
DB 366 ---GPSGQWQYAPVTGCHPVRSCITLHEDGFSAFVVAHETGHLVMEHGGQGNRCGE 422
QY 489 KSEGNIMSPTLAGNNGVFSWPCSRQVYLHFLSTAQAICLADQPKVKEYKPEKLPGE 548
DB 423 VRLGSIWAPLQAAHFRHWRSCQQLSRYLHSYD--CURDDPF--THDWPALPOLFGLH 479
QY 549 YDANTQCKWQFGEKAKLMLDFKKDICKALWCHIRGRK--CETKFWPAAEGTICGHDMWC 606
DB 480 YSMNEQCRFDGFLGYMMCTAFRTDPCQLWCSDHPDNPYFCKTKGPPDGTWCARGKHC 539
QY 607 RGQCQVYGDGPKPTHGHSWSSWPCSRCTCGGVSHRSRLCTNPKPSHGKFCGSGT 666
DB 540 FKHCIMLTLDILK-RDGNMGAWSPFGSCSRCTGTVKFRTRQCDNPHANGGRTCSGLA 598
QY 667 RTLKLNSQKCPDRSDVDFRAAOCA-----EHSRRFRGRHYKWKPYTQVEDDQDLCKLYC 720
DB 599 YDFQLNSQCPALADFREQCQWDLFEHDAQ-----HHLVPEHHRDAKERCHLYC 653
QY 721 IAREGDFFFSLSNKVKDGTPTCS-EDSRNVICDIGICERVGCDNVLGSDAEDVCGVCGN 779
DB 654 ESKETGEVSMKRVHWDGTRCSYKDAFSLVCRGDCRKGVDGIVGSKQEDKCGVCGDN 713
QY 780 SACTIIRGLYTKHHNTQYVHTIPSGARSIRIYEMNVSTSYISVRN-ALRYYLNGHW 838
DB 714 SHCKVVKGTFSRCKLGYIKMFEIPAGARHLIQEADTTSHLAVKNLETGKFIENEEN 773
QY 839 TVDWPGRYKF--SGTTFDVRSSNEPENLIATPTNETLIVELLFQGRNPGVAWEYS- 894
DB 774 DVD-FNSKFIANGVENEY-RDSDGRETQTMGLHGTITVLVIPEGDARISLTLYKYMTH 831
QY 895 -PRLGTGKQPPAQF---SYTMAIVR-SEGSVSCGGG 925
DB 832 EDSLNVDDNNVLEDDSVGEWALKKWSPCSKPCGGG 867

RESULT 4

T00355
hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura,
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The comple
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <ISH>
A:Cross-references: EMBL:AB014588; NID:G3327189; PIDN:BAA31663.1; PID:G3327190
A:Experimental source: brain
C:Genetics:
A:Gene: KIAA0688
C:Superfamily: thrombospondin type 1 repeat homology
F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>
Query Match 16.4%; Score 856.5; DB 2; Length 837;
Best Local Similarity 35.0%; Pred. No. 7.7e-54;
Matches 219; Conservative 67; Mismatches 240; Indels 99; Gaps 24;
QY 371 VETLVVVKKMQNHGHEITTVTLILNMVSALFKD----- 407
DB 220 VETLVVADDKMAAFHG-AGLKRYLLTMAAAKAFKHPISIRNPVSLVTRLVILSGGEG 278
QY 408 -----GLMKDKGTRHDH----- 445
DB 279 PQVGPAAQTILRSFCANQGLNTPEDSDPHPTALIFTRQDLCGVST--CDTLGMADVG 336
QY 446 GMSKYRSTINEDTGLGLAFTIAHSGHNFQMIHDEGNCMKSEB-----NIMSPTL 499
DB 337 TVCDPARSCAIVEDDGLQSAFTAHELGHVFNWLD-NSKPCISLNGPLSTSRHVAPVM 395
QY 500 AGNNGVFSWPCSRQVYLHFLSTAQAICLADQPKVKEYKPEKLPGEYLDANTQCKWOF 559
DB 396 AHVDPEPSPCSARFITDPLONGYGHCLLD--KPEAPLHLPVTFPGKDYADROCCQLTF 453
QY 560 GEKAKLMLDFKKDICKALWC--HRIGR-KCETKFWPAAEGTICGHDMCRGQCQVYK- 615
DB 454 GPDSRHC--POLPPPCAAALWCSDHNGHMCQTKHSPWADGTPCGPAQCMGRCJLHMDQ 511
QY 616 -DEGPKPTHGHSWSSWPCSRCTCGGVSHRSRLCTNPKPSHGKFCGSGSTRTLKLS 674
DB 512 LQDFNIPQAGGMPGMPGDCSRCTCGGVQVFSRDCRTPVPRNGKVCGRTRFRSCWT 571
QY 675 QKCPRDS-VDFRAAOCAEHSR-----RFRGRHYKWKPYTQVEDDQDLCKLYCIAEGD 727
DB 572 EDCPTGSALTFRFEEQCAAYNHRDTLFSKSPFG-PMDWVPRYTGVAPODQCKLTCQARALCY 630
QY 728 FFSLSNKKVKGDTPTCSDESRNVICDIGICERVGCDNVLGSDAEDVCGVCGNNSACTIHRG 787
DB 631 YVLEPRVVDGTPCSPDSSVVCQGRCTHAGCDRIIGSKKFKDKWCWCGDGGGSGKSG 690
QY 789 LVTKHHTNQYVHTIPSGARSIRIYEMNVSTSYISVRNLR-----RYLLNGHWT---- 839
DB 691 SPRKFRY-GYNNVVTIPAGATHILVRQOG-NFGHRSIYLALKLPDGSVALNGEYTLNPS 747
QY 840 ---VDMPG--RYKFGTTFDVRSSNEPENLIATPTNETLIVELLFQGRNPGVAWEYS- 893
DB 748 FTDVVLPGAVSLRYSGAT-----AASETLISGHGLAQPLTLQVLVAGNPQDTRLYSF 800
QY 894 -MRLGTGKQPPAQPSYTWAIVRSE 917
DB 801 FVPR-PTSTPRPTQ-DWLHRRQA 823

RESULT 5

T47158
hypothetical protein DKFzp762C1110.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47158
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24379
A:Accession: T47158
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-550 <AAA>
A:Cross-references: EMBL,AL162080
A:Experimental source: adult melanoma (Mewo cell line); clone DKFzp762C1110
C:Genetics:
A:Note: DKFzp762C1110.1

Query Match	14.0%;	Score	731.5;	DB	2;	Length	550;
Best Local Similarity	36.0%;	pred.	No. 4.9e-45;				
Matches	164;	Conservative	61;	Mismatches	197;	Indels	33;
Gaps	15;						
Qy	491	EGNIMSP	TLACNGVFSWSP	CSRQYLHKLFLSTAQAI	CLADOPKVPKVEKYPKELF	GLYD	550
Db	9	DSHMA	SNLNLDSQSP	PCSAAYMIT	SFLDNHGCE	LMD--KPNPIQL	EGDLP
							PGTSYD
Qy	551	ANTQCKW	OFGEKAKL	CMLDPKDKICKAL	WCHRGRK-----CEIKF	MPAAEGTIC	GHDMW
Db	67	ANRO	CCQFTFGED	SKEC--PDAAST	CSLWLC--TGTSG	GVLYCQTHFF	FWAGTSC
							GBGW
Qy	606	CRGGCV	KYGBEG--PKPTH	GHWSDSWS	PCSRTCGGV	SHRSRLCTNPKP	SHGGK
Db	123	CINGK	CVNKTDRKH	EDTTFPHSG	GNWGPWGDCSR	TCCGGVQVYTW	RECDNPV
							KNGK
Qy	664	GSTR	TLKLCNSOK	PCP-RDSV	OPRAAQAEHN-----SRRFR	GRHYWKVP-YTQ	VEDQDLCK
Db	183	GKRV	YRSCNLEDC	PDNNKGTFR	EEQCAHNEF	SKASFGSP	FAEWIPK
							YAGVSP
Qy	718	LYCIA	EGDFPFFSL	SNKVKGDT	PCSEDSRNV	CDIGICER	VGCDNLV
Db	243	LICQ	AKGIGYFVL	QPKVVDTP	PCSPDSTS	VCVQCVKAG	CDRIIDS
							KKKFKD
Qy	778	NNSACT	TIHRGLYTK	HHHNNQYH	VMVTIISG	ARSIRIYEMNV-----STSY	ISVRNAL
Db	303	NGST	CKKISGS	SVTS--AKPG	VHDIIITITPG	ATNIEVKQRN	QGRSNNGS
							FLAIIKA
Qy	833	VING	HTVTDW	PCGR-YKFS	GTTFDYR	SRSENEPNLI	ATGPTNET
Db	361	ILNG	YTTIST	LEQD	IMYKGV	LVRYGSSA	ALERIS
							FSPLKE
Qy	890	WEYS	MPRLGTE	KOPPAQ	PSYTWAI	VR-SEC	SVSC
Db	421	YTYF	YVKK--KKES	FNALPT	FSAMVIE	WEGEC	SKSC
							453

RESULT 6
 T18856
 angiogenesis inhibitor homolog - Caenorhabditis elegans
 C.Species: Caenorhabditis elegans
 C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C.Accession: T18856; T24653
 R.McMurray, A.
 submitted to the EMBL Data Library, July 1995
 A.Reference number: Z19031
 A.Accession: T18856
 A.Status: preliminary; translated from GE/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-1444 <Wtl>
 A.Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1
 A.Experimental source: clone C02B4
 R.McMurray, A.
 submitted to the EMBL Data Library, July 1995
 A.Reference number: Z19917
 A.Accession: T24653

A>Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-1444 <MI2>

A:Cross-references: EMBL:Z50006; PDB:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1

A:Experimental source: clone T07C5

C:Genetics:

A:Gene: CESP:C02B4.1

A:Map position: X

A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 510/3

Query Match 10.0%; Score 525.5; DB 2; Length 1444;
Best local similarity 22.4%; Pred. No. 1.4e-29;
Matches 213; Conservative 130; Mismatches 313; Indels 293; Caps 44;

QY 123 HVLQYDLVSAVEYDHRGDYVSHEIMHQRRAVAVESVSLHLRLK--GPRDHFMDL 180
DB 7 HHLNEELKQVFGSNKHDVPEYSIEATRH-----PLKGNLKMKTANNDTYHLNL 59
QY 181 RTSASLVAPGPIVTLGKTGKSVQT--LPPEDFCFYGSLSRHSRNSPSHGKFCGEGST 237
DB 60 RKNRIVSPHII--SVVRHGDDVTYAGLRDVEQCHYQGEVYKSHGNMKA----- 107
QY 238 RTLKLCNS-----QKPROSVDFRAQCAEHNRRPRGRYKWKPYTQVEA 283
DB 108 -AISDCGALMGSI VMEDHFLVLQTLPK-----RVHHLQKRLHYKRSAGLUTNA 156
QY 284 DLCKLYCIAEGDFPFSLNSKNKVG--TPCSEDSRNVCIDGICELS-----V 328
DB 157 E-----SKIREITRLQEQSEFC--DTSEQLDDPAMTIPAHLEFNYT 197
QY 329 VSTSAHPQPPKEDFILPDEYKSLRHKRSLRSHRNEELAVETLVVVDDKQMQNHGHE 388
DB 198 IPTSAQL---DSSSIFIFN-----MDPITLIEIGFLDLSKLFEPERE 235
QY 389 -----NITVYLTILNMVSLFXDGLM----- 410
DB 236 YIQDAEQHLLEFSLALNNVLYQODTLPNLDIVIVRYEHWRTOPSLSTGVHKNQGA 295
QY 411 -----GKDGTL--RHDHAILTLGLDSCWKNPCDPLGFAPISGMSKY 451
DB 296 QSLLDADFQRYQAHHNPGDITLDMNHVHGVLITGYDI---YHTTTSVAGVAPVARMCDPL 352
QY 452 RSCITNETGLCLATIAHESGHNFMGHDGEGNMCKSEGMISFTLAGRNGVFWSPC 511
DB 353 FACSLVEGLHLGRSFVLAHEMGNHGMVHDGVQNCNK--GGLMSAVANGA--GKTTWSDC 409
QY 512 SROYLHKEL---STAQATCLAD--QPKPKYKYPE--KLPGELYDANTOCKWOFGEKAKL 565
DB 410 SVREFNAPLLQDSEGRGNCLDAPGLISTNHLSDLRLPQGFADQCCSYFWRDVKV 469
QY 566 CMLDPK--KDICKALWCHRIGRKCTKFWPAEAGTICGHDMWCRGGQCVKYG--DEGPKPT 622
DB 470 EIPNGKAMDDICRILWCGNSGSGTISTAH--PALEGSCGANKWKCHGQCTHTFGLTPVPI 528
QY 623 HGHWSDW---SSWSPCSRTCGGG---VSHRSILCTNPKPSHGKFCGEGSTRTKLNSQK 676
DB 529 DGESEWGAEGKCFIQQCAVSGSITVQQHHRDCVNPAPNGKCTCEGANIRIVCGA-- 586
QY 677 CFRDSVDFRAQCAEHNRRPRGR---HYKWKPYTQVEDQDL-----CKLYCIA 722
DB 587 -----TSSNCLGFTREFGNKICSSIKYDPHK--PDQQLTGEGFEHSTQPCRVMCHL 636
QY 723 EGDFPFSLSNK--VKDGTPCSEDSRNVCIDGICERVGCDNVLGSDAEDVCGVCNNGN-- 779
DB 637 IGSEL---IRNKQGFDPDGTGCFDA--YCVGQCALSCDNLKALVBPQED--CPRIEGRSV 690
QY 780 -----SACTIHRGLVTKHHHTNQYVHMVTFPSGARSIRIYE-----MNVSTSYI 823
DB 691 HGWBSWSWSECSVCSGL-----GGREVERKCSSGRKCGQVSESR 732
QY 824 SVRNALRYYLNGHWTVDMPGKYKSGTTFDVSRYNEPENLIATGFTNE--TLIVLELFF 881
DB 733 PCEGLVLRDCEEFGEWK--EW-----GSCSEKALGVQKRF 765

QY 882 QGRNPGVAWEYSMPRLGTGKPPQPSVT----WAIVRSECVSCGGGR 926
Db 766 R---PCLTDQSSXHLQBERPCDNEGCTWWDEN-----SSCSQCGGR 807

RESULT 7
C89114
Protein C37C3.6a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89114
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
A:Accession: C89114
A:Status: preliminary
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1558 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a
C:Genetics:
A:Map position: 5

Query Match 9.7%; Score 508.5; DB 2; Length 1558;
Best Local Similarity 33.6%; Pred. No. 2.6e-28;
Matches 110; Conservative 58; Mismatches 134; Indels 25; Gaps 9;

QY 602 HDMCRGQCVKYVDEGPKPTHGWSWSSPCSRCTCGGVSRRSLCTNPKPSHGKGF 661
Db 56 YQYVVDGVSVTVDKSGQKET--GNWGPWVPENECSRSCGGVQLEKRCQSGD----- 106
QY 662 CEGSTRTLKLCNSQKCPDRSDVDFRAAQAACAEHNSRRFRGRHYKWKPYTQVEDQDLCKLYCI 721
Db 107 CTGASVRYISCNLNAAC--ESGTFDFAEQCSKFNDALDGNVHKWTFY---RGNKNCCELVCK 162
QY 722 AEGFOFFSLSNKVKVKGCTPCSEDSRNVCIDGICERVCNDVLSGDAVEDVCGVCNGNNSA 781
Db 163 PESGNFYKWKADKVDGTFKCDKSKNDICVDGECLPVGCDKGLSSLFKDKGCKDGDGST 222
QY 782 CTIHRGLYTKHHHTNQYHYHMTIPSGARSIRIYEMNVSTSYISVRNALRYYLNGHWTVD 841
Db 223 CKTIEGRDERNLSPGYHDIILKPEGATNIKIQAARKSTNNLALKNGSDHFLYLNGLIQ 282
QY 842 WPGRYKFSGTTFDYRRSYNEPENLIATGPTNETLIVELLFO--GRNPGVAWEYSMPRLGT 899
Db 283 VEKEVEVGTTIFVYDDA--EPETLSAQGPLSEELTVALFRKGRSDRTAIRKYEFSIP---L 337
QY 900 EKQPPAQPSY-TWAIVRSECVSCGGG 925
Db 338 EEEVDVYMKFDNW----TPCSVSCGKG 360

RESULT 8
T34395
hypothetical protein C37C3.6b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000
C:Accession: T34395; T34394
R:Geisel, C.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid C37C3.
A:Reference number: Z21518
A:Accession: T34395
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2167 <GEI>
A:Cross-references: EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C37C3.6b
A:Experimental source: strain Bristol N2; clone C37C3
A:Accession: T34394

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1555, 'SKF', <GE2>
A:Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a
A:Experimental source: strain Bristol N2; clone C37C3
C:Genetics:
A:Gene: CESP:C37C3.6b; CESP:C37C3.6a
A:Map position: 5
A:Introns: 32/3; 104/2; 156/2; 207/1; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 1555/1

Query Match 9.7%; Score 508.5; DB 2; Length 2167;
Best Local Similarity 33.6%; Pred. No. 3.9e-28;
Matches 110; Conservative 58; Mismatches 134; Indels 25; Gaps 9;

QY 602 HDMCRGQCVKYVDEGPKPTHGWSWSSPCSRCTCGGVSRRSLCTNPKPSHGKGF 661
Db 56 YQYVVDGVSVTVDKSGQKET--GNWGPWVPENECSRSCGGVQLEKRCQSGD----- 106
QY 662 CEGSTRTLKLCNSQKCPDRSDVDFRAAQAACAEHNSRRFRGRHYKWKPYTQVEDQDLCKLYCI 721
Db 107 CTGASVRYISCNLNAAC--ESGTFDFAEQCSKFNDALDGNVHKWTFY---RGNKNCCELVCK 162
QY 722 AEGFOFFSLSNKVKVKGCTPCSEDSRNVCIDGICERVCNDVLSGDAVEDVCGVCNGNNSA 781
Db 163 PESGNFYKWKADKVDGTFKCDKSKNDICVDGECLPVGCDKGLSSLFKDKGCKDGDGST 222
QY 782 CTIHRGLYTKHHHTNQYHYHMTIPSGARSIRIYEMNVSTSYISVRNALRYYLNGHWTVD 841
Db 223 CKTIEGRDERNLSPGYHDIILKPEGATNIKIQAARKSTNNLALKNGSDHFLYLNGLIQ 282
QY 842 WPGRYKFSGTTFDYRRSYNEPENLIATGPTNETLIVELLFO--GRNPGVAWEYSMPRLGT 899
Db 283 VEKEVEVGTTIFVYDDA--EPETLSAQGPLSEELTVALFRKGRSDRTAIRKYEFSIP---L 337
QY 900 EKQPPAQPSY-TWAIVRSECVSCGGG 925
Db 338 EEEVDVYMKFDNW----TPCSVSCGKG 360

RESULT 9
T00260
hypothetical protein KIAA0605 - human
C:Species: Homo sapiens (man)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00260
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Chara, O.
DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The complete coding sequences of 100 genes, including KIAA0605, were determined from cDNA clones.
A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00260
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-951 <NAG>
A:Cross-references: EMBL:AB011177; NID:g3043733; PIDN:BA25531.1; PID:g3043734
A:Experimental source: brain
C:Genetics:
A:Note: KIAA0605
C:Superfamily: thrombospondin type 1 repeat homology
F:46-106/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 9.5%; Score 499.5; DB 2; Length 951;
Best Local Similarity 33.9%; Pred. No. 6.1e-28;
Matches 111; Conservative 51; Mismatches 134; Indels 31; Gaps 9;

QY 609 GQCVKYVDEGPKPTHG-----WSDWSWSPCSRCTCGGVSRRSLCTNPK--- 654
Db 22 GDTVSTGSTNSPTNSLEGGTATAPWGEWTKWTFASRSCGGVTSQERHCLQQRKS 81
QY 655 -PSHGKFCBEGSTRTLKLCNSQKCPDRSDVDFRAAQAACAEHNSRRFRGRHYKWKPYTQVEDQDLCKLYCI 708
Db 82 VPFGNRTCTGTSKRYQLCRVQCPCPDGSRFSRQCVCSFNSHYNGRTHQMKPLPYDDYV 141
QY 709 QVEDQDLCKLYCIAEGDFPFSSLSNKVKGTGTPCS-EDSRNVCIDGICERVCNDVLSGDA 767

Db 142 HISSRP-CULHCTT---VQQRQLMVPARDGTSCKUTLURGVCSVKCEPIGCDGVLFSTH 198
QY 768 VEDVCGVGNNSACTIHRGLYTKRHHHTNQYHYHMTIPSGARSIRIYEMVSTSYISVRN 827
Db 199 TLDKCGICQGGSSCTHTVGNRYKGNVHLSVTHIPAGARDIQIIVERKKSADVLALAD 258
QY 828 ALRRYLLNCHWTDPGKFKSGTTFDYRRSVNEPEN-----LIATGPTNETLIVLELFP- 882
Db 259 EAGYFFNGYKVDSPKFNFIAGTVKVRPMVYETGIEYIAQSGPTNOGLNVMVWQON 318
QY 883 GRNPGVAVESMRLGTEKOPPAQPSY 909
Db 319 GKSPSITFEYTLLOPPHESRP--QPIY 343

RESULT 10

Tl6892
hypothetical protein Tl6892.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: Tl6892
R:Bentley, D.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid Tl6892.
A:Reference number: Z18599
A:Accession: Tl6892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-860 <BEN>
A:Cross-references: EMBL:U42846; NID:g1125809; PID:g1125810; PIDN:AAA83600.1; CESP:Tl9D2
C:Genetics:
A:Gene: CESP:Tl9D2.1
A:Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2; 505/3; 58

Query Match 9.0%; Score 472.5; DB 2; Length 860;
Best Local Similarity 25.3%; Pred. No. 4.8e-26;
Matches 187; Conservative 72; Mismatches 241; Indels 239; Gaps 36;
QY 364 HRNEELNVETLVVD-----KKMQNHGHENITTVTLNVMVSALFKDGLM----- 410
Db 128 HRDLITVELAVFADDAMWDFKMYKAAEENHTFIMAVVNMIDVLTORLQPRINI 187
QY 411 -----GKDGTRHDHAILLGLD 427
Db 188 KIVRYEILKNPILMNAARKHSGVDRLLDAPCOYQNEINPPNDADPRHNDHALFSGYD 247
QY 428 ICSMKNEPCDTLGEAPISGMSKYRSCITINEDTGLGLAFTAHESGHNFHMGHGEQNM 487
Db 248 L--HRNGVKTVAGAPVKMGCSGVRSCITINEGLDFGVSFVVVTHMGHSLGMYHDGD-NEC 304
QY 488 KKSEGNIMSPTLAGRNGVFSWSPCSROYLHKFL-----STAQAICLAD---QPKPVKEYK 539
Db 305 -DLRCCIMSPVG--SGKTHWSQCSVNEVEMATFVGHGLDGFPPNCLQDASANEQRMVAFK 361
QY 540 YPEKLPBELDYDANTQCKQFOFEKALCMLDPK--KDIKALWCHRIGRKCTKMPAAEG 597
Db 362 ESEP-PQGLFTLDQCEIFHGECEWKHELKDGQTMQNICMVWCGN-GEGVIRTAHPALEG 419
QY 598 TICGDMWCRGQCVKYGDEGPKPTHGHWSDWSSWSPCSRTCCGGVSH-----RS 647
Db 420 TYCGFGMICRQGCVG--SSQMLRVTVGGWSTWND-RP-APTCCGRCSQCEIRGQIRIMRS 476
QY 648 -RLCTNPKPHGGKFCBEGSTRTLKNSQKCPRDSVDFRAAQCAEHNRSRFRGRHYKWK 706
Db 477 IRQCNFFSNNGGAPCCQDEARGMVCHRDVCGDSIENYATVCS-----RLRDNWAIP 530
QY 707 YTOVE-----DQDLCKLYCIAEGDFFPSLSNKKVDTGTPCSBDSRVNIDGICERVGC- 759
Db 531 NTLISGGMQFEQAMCKIWLISGNTIRTVSN-FPDGAPCGPG--QYCIKGECPRLCG 587
QY 760 -----DNVLG-----SD 766

Db 588 STTLAYSEADCPLSVLQTTTPMPHVHSDVQFAGKTNPYKEHKHTTFLNEWSQMSWSE 647
QY 767 AVEDVC-----GVCNG---NNSACT-----IHRGLYTK-HHHTNOYHYMVT 803
Db 648 CVTYDCHTQGVKVRVRCLAGVACALRERQPCTRPCTGSEPLTTTSPFQCTFNRRIAP 707
QY 804 IPSGARSIRIYEMVSTSYISVRNALRRYYLNGHWTWDFGRYKFSGTTTFDYRRSYNEPE 863
Db 708 LP-----NRQTNMI-----LRKV---DHMG-PWSACSVTCTGGOKLRRR---E 743
QY 864 NLIA-----TGPTNETLIVLELFGGRNPGVAVESMRLGTEKOPPAQPSYTWAIIVSE 918
Db 744 NCIGQECAGTGP-----CMQSCRENKNTWTWS-----QW-----SDC 777
QY 919 SVSCGGGRCLPVLLEAAC 937
Db 778 SVNCGEG-----VQPRKRA 792

RESULT 11

Tl5976
hypothetical protein F08C6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: Tl5976
R:Bentley, D.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F08C6.
A:Reference number: Z18440
A:Accession: Tl5976
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-957 <BEN>
A:Cross-references: EMBL:U99378; NID:g868184; PID:g868185; PIDN:AAA68721.1; CESP:F08C6.
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F08C6.1
A:Introns: 23/1; 135/3; 220/3; 285/3; 325/2; 376/3; 461/2; 498/3; 532/3; 627/2; 714/1;

Query Match 5.9%; Score 310; DB 2; Length 957;
Best Local Similarity 20.8%; Pred. No. 3.1e-14;
Matches 172; Conservative 73; Mismatches 225; Indels 356; Gaps 41;
QY 263 HNSRRF--SGRHYKWKPYQVEADLCKLYCIAEGDFFFSLSNKKVKGDTGTPCSDSRVCI 320
Db 157 HRSKRATENGAVVYKRET-----LAGEPKD----- 182
QY 321 DGICELSVYSTSAHMPQPK--EDLFLPDEYKSLRHKRSLRSHRNEELNVETLVVV 377
Db 183 --FCGLDNVVTEESLVEDESAIFEDVFTGQ-----RLTQSDLLIVELAVFV 227
QY 378 DKQMQN---HG---HENITTVTLNVMVSALFKD----- 407
Db 228 DENLMHFHSSKSGMADRKLQDYTLTLNNIQIMYQPTASPLTFRVIRYEVLTQPSA 287
QY 408 --GLMKCKGTRH-----DRAILLTGLDICSWKNEPCDTLGFAPISCM 447
Db 288 LAGYLHNHGNQMYLDRFCYQRYNLAVRDWDHAILMTGYDIHRGASR--SISGIARLDGM 346
QY 448 CSKYRSCITINEDTGLGLAFTIAHESGHNFHMGHGEQNMCKSEGNIMSPTLAGRNGVFS 507
Db 347 CDPWNCTTLEAGLDGFTSAFIGTHGHRFTLKD----- 380
QY 508 WSPCSROYLHKFLSTAQAICLADQPKPVKEYKYPKLPGLDYDANTQCKMQFGKAKLCM 567
Db 381 ----- 380
QY 568 LDFKXDKICALWCHRIGRKCTKMPAAEGTTCGHDWRCRGQCVKY--GDEGPKPT----- 622
Db 381 -----HKSDTLF-----TGTPCGPSKWCQLGRCVFWTGTNBIQPTVQHV 419
QY 623 -----HGHWSMSSWSPCSR--TCGG-----GVSHSRSLCTNPKPHSGGKFCBGS 665

Db 420 APVVTTLPSRIDGWSGWA-TTCSQCTCNGILGSLVGLAIAARTCSAPYPANGSGDCVGS 478
QY 666 TRTLKLCNSQKCR--DSVD-FRAACAEH-----NSRRFRGRHYKWKPYTOVEDDOLCKL 718
Db 479 TSRAVLIC-SRQCGRASVDEXTSDCKMEQRLKNDRELTGKSQLNR-----PQACKV 533
QY 719 YCIAE-----GFDFPFLSNKVKDGTSPCESNRVVICDIGICERVGCN--VLGSD--A 767
Db 534 FCDVQHYGSQRYRFF--GDNLPDGTSGYD--RYCLDGECLANCNNALISDQSC 588
QY 768 VEDVGV-----CNG-----NNSACTIHRGLYTKHHNTQYY 799
Db 589 PTDCTPITDQSSVYRQGMWTSLWTSCTATCGGYRKNRACSI-TQCEGNEDETEVC 647
QY 800 HWYTIPEGARSIRIYENVUSTSYISV--RNALRY--YINGHWTV--DWPGRYKESG-- 850
Db 648 SSESCHSVLRVGNENSTWTEWNCSCVSGRGSQARYKCLSPHRTLAFCPGNKVYNEL 707
QY 851 -TTFDYRRSY-----NEPENLIA-----TGPTNETLIVELLFQGRNPGVAWE 891
Db 708 RITFFKARSYIMCSVRCKIKRTISEKNIEVRSCDNGPCNAI-----GV-WG 754
QY 892 YSNPRLGTEKOPPAQPSYTHAIYRSECSVSCGGRCPLVLLLEAAC 937
Db 755 -----TWG-GWSTCSTSCGPG-----TLVRQRTC 777

RESULT 12
A60385
monocyte surface antigen MS2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 26-Aug-1999
C:Accession: A60385
R:Yoshida, S.; Setoguchi, M.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.
Int. Immunol. 2: 585-591, 1990
A:Title: Molecular cloning of cDNA encoding MS2 antigen, a novel cell surface antigen s
A:Reference number: A60385; MUID:91197896; PMID:1982220
A:Accession: A60385
A:Molecule type: mRNA
A:Residues: 1-826 <YOS>
A:Cross-references: EMBL:X13335
C:Superfamily: mouse meltrin alpha; disintegrin homology
C:Keywords: glycoprotein; surface antigen; transmembrane protein
F;1-14/Domain: signal sequence #status predicted <SIG>
F;402-484/Domain: disintegrin homology <DIS>
F;659-683/Domain: transmembrane #status predicted <TM>
F;330/Active site: Glu #status predicted

Query Match 5.8%; Score 305; DB 2; Length 826;
Best Local Similarity 21.8%; Pred. No. 5.9e-14;
Matches 163; Conservative 79; Mismatches 238; Indels 274; Gaps 37;

QY 152 RRRRAVAV---SEVESLHLRLKGRHDFHMDLRTSSSLVAPGFIQVTLGKTGKSVQTLP 208
Db 39 RSRRLPSHWGQYPESLSVALGTSGHVFLHLKRNLDLGGSSYTYSAANGSEVTEQLQ 98
QY 209 PEDFCYQGLSRSHRNSPSHGKFCBGSTRTLKLCNSQK-----PRDSVDFR 256
Db 99 EQDHCLYQGHVEGY-----EGSAASISTCAGLRGFRVGVSTVHLIEPLDADE-- 145
QY 257 AAQCAEHSNRFRGRHYKWK-PYTOVEADLCKLYCIAEGDFPFSLSNKVKDGTCPSEDS 315
Db 146 -----EGQHAMYQAKHLQKAGTC-----GVKD----- 168
QY 316 RNVICDIGELSVVSTSAHMPQPPKEDLFLPDEYKSLRHKRSLRSHRNEBLNVETLV 375
Db 169 -----TWINDLGRALIEYRAQPN---WLIPETR-----YVELYV 202
QY 376 VDKKMMQNHG-HENITVTVTLINMVSALFKD----- 407
Db 203 VADSQBFQKLSREAVRQVRVLEVVNVHVDKLYQELSPRVVLVGLIWNKDKFYISRYANVT 262
QY 408 -----GLMGKDGTRHDHAILLTGLDICSWKNEPCDTLGFAPISGMCSYRSCCTIN 457

Db 263 LENFLSWREQLQGQH--PHDNVQLITGVDFIG-----STVGLAKVSALCSRH-SGAVN 313
QY 458 ED---TGIGLARTIAHESGHNFEMTHDGE--GNMC---KKSGNIMSPFLAGNGVFSHS 509
Db 314 QDHSKNSIGVASTMAHELGNLGMHSHDEIPQCYCPEPREGGCGCINTESI-GSKFPRIFS 372
QY 510 PCSROYLHKFLSTAQAIADLADOPK-----PVKEYKYKPEKLPGELYDANTCKWQFGSK 562
Db 373 RCKSIDLESFVTKPOTGCLTNVPDVRNRFVGGPVCGNLFVEH--GEOCDGCTP----- 422
QY 563 AKLCMLDFKKDKTCKALWCHRI-GRKC-----ETKFMFAAE----- 596
Db 423 -----QDCQNPCCNATTQOLVKGACASCTCCECKVKPAGVCRLSKDKDLEEFCDGR 477
QY 597 GTICGHDMV-----CRGQCQVKYGDGPKPHGHWSWSSWSPSRT----- 638
Db 478 KFTCPEDAFQONGTFCPGGYCF-----DGSCTFLAQQCR-DLWGPAGARVAADSCYTFSDP 532
QY 639 -CGGGV-SHRSLCTNPKPSHGKFCBGSTRTLKLCNSQKCPRDSVDFRAACCAHNSRR 696
Db 533 GNGRMYSGRINRC-----GALYCEGGQKPL-----ERSFCT 564
QY 697 FRGRHYKWKPYTOVEDDQLCKLYCIAEGDFPFSLSNKVKDGTCPSESRNVICIGICR 756
Db 565 FSSNH-----GVCHALGTGSNIDTF-----ELVLOGTKCEEG--KVMGDSGCD 606
QY 757 VGCDNVLGSDAVEDVC---GVCNGNNSACTIHRG 787
Db 607 L---RVYRSEKSAKCNHNGVCN-HKRECHCHKG 636

RESULT 13
T22836
hypothetical protein F57B7.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22836
R:Lennard, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19623
A:Accession: T22836
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-508 <WIL>
A:Cross-references: EMBL:Z74037; PIDN:CAA98493.1; GSPDB:GN00023; CESP:F57B7.4
A:Experimental source: clone F57B7
C:Genetics:
A:Gene: CESP:F57B7.4
A:Map position: 5
A:Introns: 45/3; 137/2; 221/2; 256/2; 306/2; 409/3; 451/3

Query Match 5.5%; Score 289; DB 2; Length 508;
Best Local Similarity 23.8%; Pred. No. 4.6e-13;
Matches 129; Conservative 70; Mismatches 185; Indels 158; Gaps 24;

QY 198 KTGTSVOTLPEDFCFYQGLSRSHRNSPSHGKFCBGSTRTLKLCNSQKCPRDSVDFRA 257
Db 65 KTQELSLDLVIVADFLSYQAFLM-SNGDSHR----- 95
QY 258 AACAEHSNRFRGRHYKWKPYTOVEADLCKLYCIAEGDFPFSLSNKVKDGTCPSEDSRN 317
Db 96 ---A1HN-----LKEYLHALFEQTKI--IYDGISFGNETLHMVFAGTWIATQERD 140
QY 318 VICIDIGELSVVSTSAHMPQPPKEDLFLPDEYKSLRHKRSL-----LASHRNE 367
Db 141 -----CPLWI-----SWAESEERVNLNEIRLEEKRODLSNFTFVDDTFMNSTDSD 187
QY 368 ELNVTETLVVD--KKMMQ-----NHGHENITVTVTLINMVSALFKDGLMGKQCT--- 416
Db 188 NSSTDALLSSDMPKLRKRFVDITLSEMQENSTWTLKIDSKAIDKKTIMLKQGTGLPR 247
QY 417 HDHAILLTGLDICSWKNEPCDTLGFAPISGMCSYRSCCTINEDTGLGL-AFTIAHESGHN 475

Db 248 HEHAVLITFDLIS-INGNSATCGMAYVGNICENGDSVSSWEDIGAGLTSILMAHEIGHS 306
 QY 476 FGMHIDG--EGNMCKKSEGNISPTUAGRNQVFSM-----SPCSRQYLHKFLSTAQATC 527
 Db 307 LGALHDGAYETAECDSNDNYLMAVAVSGADRQSFNLNRRMNCNSINSIENLKPSFTANC 366
 QY 528 LADQPKVKEYKYPE-----KLPGELVDANTQCKWQGEKAKLCM---LQPKDI 574
 Db 367 -----VKKWKTKGKOVQSKDFIKKPGELVKITROCCQVAFGPTFIPCLHIGYFHEQSI 419
 QY 575 CKALWC-HRIRKCT-KMPAAEGTRICGHDMWRCGGQCVKYGDEGPKXPHGHWSWSW 632
 Db 420 CERINCSGDESECQTLNFPADGTGECYNNMWCLEGSVCQ-----NTKX 465
 QY 633 SPG-----SRTCGGVSHRSLCTNPKPSHGKFCGEGSTRTLK-LNSQKCPRDSVDFRAA 687
 Db 466 MCKDINSKTC-----KYSTSKLHYCKSK-----DFREI 496
 QY 688 QC 689
 Db 497 CC 498

RESULT 14

A55796

ecarin precursor - saw-scaled viper

C:Species: Echis carinatus (saw-scaled viper)

C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 21-Jul-2000

C:Accession: A55796

R:Nishida, S.; Fujita, T.; Kohno, N.; Atoda, H.; Morita, T.; Takeya, H.; Kido, I.; Paine

Biochemistry 34, 1771-1778, 1995

A:Title: cDNA cloning and deduced amino acid sequence of prothrombin activator (ecarin)

A:Reference number: A55796; MUID:95151760; PMID:7849037

A:Accession: A55796

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-616 <NIS>

A:Cross-references: GB:D32212; NID:G717090; PIDN:BA06910.1; PID:G717091

C:Superfamily: mouse meltrin alpha; disintegrin homology

F:404-486/Domain: disintegrin homology <DIS>

F:338/Active site: Glu #status predicted

Query Match 5.1%; Score 269.5; DB 2; Length 616;
 Best Local Similarity 21.0%; Pred. No. 1.5e-11;
 Matches 134; Conservative 88; Mismatches 210; Indels 205; Gaps 33;

QY 202 KSVQTLPP-EDFCFYQGLSLRSHRNSPSHGKFCGEGSTRTLKLCNSQKCPRDSVDFRAAC 260
 Db 94 REITTPSVEDHCYHGRIONDAB-----STASISACNLK----- 129
 QY 261 AEHNSRRFRGRHYKWKPYQVEADLCKLYCIAEGDFDFFSLSNKVKDGTPTCSDESRNVCI 320
 Db 130 ---GHFKLGETYFIEPLKTPDSEAHAVY-----KYENIENE-----DEAPKMC- 170
 QY 321 DGICELSVSTSAHMPQPKEDL-FILPDEYKSLRHKSLRSHRNEELNVTLVVVDK 379
 Db 171 -----GVTQDNWSEDEPIKTLGLVPP-----HERKF-----EKKFIELVVVDH 211
 QY 380 KMQNHGHEH-ITTVTLTNMVSALF-----KDLGMLG-----KDGTL- 415
 Db 212 SMVTKYNNDSATRTIYEMLNTVNEIYLPFNIRVALVGLFECNGLINVTSTADTLH 271
 QY 416 -----RHDHAILLTGLDICSWKNEPCDTLGFAPISGCMCKYRSCCTINEDTG 461
 Db 272 SPGEWRASDLLNKRHDHQAQLLNVTL-----DHSTLGTITFYVYCKSDRSVELLIDYS 325
 QY 462 ---LGLAFTAHESGHNFHMGEGNMCKKSEGNISPTUAGRNQV----FSWSPCSRQY 515
 Db 326 NITFNMAIITAHESGHSLGMLHDTKFTC-----GAKPCIMFGKESIPPPKPFSSCSYDQ 380
 QY 516 LHKFLSTAQAICLADQPKVKEYKYPEKLPGLGYDANTQCKWQGEKAKLCMLDFKDKLIC 575

Db 381 YNKVLLKYNPKCILDPPLR-KDIASPAVCGNEIWESEEC--DCGSPA-----DCRNPPCC 432
 QY 576 KALACHRTGRKCEYKFMPPAAGTTCGHDMMC-----RGQCCKYKGGEGKPHGHWS 628
 Db 433 DAATC-----KLKPGAE---CGNGECDCDKIRKAGTECPARDDCDVAEH----- 475
 QY 629 WSSNSPCSRCTCGGVSHRSLCTNPKPSHGKFCGEGSTRTLKLCNSQKCP----- 678
 Db 476 -----CTG-----QSAECPRNEFORNQPCUNNS---GYCVNGOCPIMLNQIALF 518
 QY 679 RDSVDFRAAQCAEHSR-----RFRGRHYKWKPYQVEADLCKLYCIAEGF-- 725
 Db 519 SPSATVAQDSQFQRLQSSYYGYCTKEIYGYKRPFCAP-----QDVKGRLVCLDINSFKK 574
 QY 726 -----DFFFLSNK--VKDGTPTCSDESRNVCIQDGC 754
 Db 575 NMRCKNDYSADENKNGIPEPGTKC-EDGK-VCINRKC 609

RESULT 15

S48160

metalloproteinase (EC 3.4.24.-) H-I precursor - carpet viper

N:Contains: disintegrin

C:Species: Echis pyramidum leakeyi

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000

C:Accession: S48160

R:Paine, M.J.I.; Moura-Da-Silva, A.M.; Theakston, R.D.G.; Crampton, J.M.

Eur. J. Biochem. 224, 483-488, 1994

A:Title: Cloning of metalloproteinase genes in the carpet viper (Echis pyramidum leakeyi)

A:Reference number: S48160; MUID:95010025; PMID:7925363

A:Accession: S48160

A:Molecule type: mRNA

A:Residues: 1-617 <PAI>

A:Cross-references: GB:X78970; NID:G763092; PIDN:CAA55565.1; PID:G763093

C:Superfamily: mouse meltrin alpha; disintegrin homology

C:Keywords: hydrolase; metalloproteinase; venom

F:1-18/Domain: signal sequence #status predicted <SIG>

F:193-617/Domain: propeptide #status predicted <PRO>

F:193-617/Product: metalloproteinase H-I #status predicted <MAT>

F:403-485/Domain: disintegrin homology <DIS>

F:337/Active site: Glu #status predicted

Query Match 5.0%; Score 263.5; DB 2; Length 617;
 Best Local Similarity 18.9%; Pred. No. 4.1e-11;
 Matches 144; Conservative 90; Mismatches 206; Indels 323; Gaps 33;

QY 202 KSVQTLPP-EDFCFYQGLSLRSHRNSPSHGKFCGEGSTRTLKLCNSQKCPRDSVDFRAAC 260
 Db 94 REITTPSVEDHCYHGRIONDAB-----STASMSACNLK----- 129
 QY 261 AEHNSRRFRGRHYKWKPYQVEADLCKLYCIAEGDFDFFSLSNKVKDGTPTCSDESRNVCI 320
 Db 130 ---GYFMLRGETYFIEPLKIPDSEAHAVY-----KYENVEKE-----DEAPKMC- 170
 QY 321 DGICELSVSTSAHMPQPKEDL-FILPDEYKSLRHKSLRSHRNEELNVTLVVVDK 380
 Db 171 -----GVTQTNWSEDELKASQLVATSE-----QQRSYKYIEFVTVVADY 211
 QY 381 VMQNHGHEH-ITTVTLTNMVSALF-----KDLGMLGDT----- 415
 Db 212 MYRTKYNNDSATVRRIYEVINLNMVYIVFNHVALTHIEIWNSTHDQITVQSAADTLDL 271
 QY 416 -----RHDHAILLTGLDICSWKNEPCDTLGFAPISGCMCKYRSCCTINED- 459
 Db 272 FQDMRAKULLTKKHNDNAQLFTGINLNG-----QTLGIARMGCMCSNSVGVLDYCK 325
 QY 460 TGLAFTAHESGHNFHMGEGNMCKKSEGNISPTUAGRNQVFSWSPCSRQYLHKF 519
 Db 326 NYLLVAITWAHELGNLGMHDH-NGN-CNCPDTSICMSAVAGPEPVFSFNSCNRNDYRSF 383
 QY 520 LSTAQAICLADQPKVKEYKYPEKLPGLGYDANTQCKWQGEKAKLCMLDKKICKALW 579
 Db 384 RNSQSKCIDNKP-----LKTDIVSPSV 406

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 21:53:59 ; Search time 48 Seconds
(without alignments)
932.696 Million cell updates/sec

Title: US-09-981-151A-8

Perfect score: 5236

Sequence: 1 MKPRARGWRGLAALWMLLAQ.....LEAACQPSATYALAPLES 952

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2408	46.0	1081	1 AT18 HUMAN	Q8te60 homo sapien
2	1223	23.4	1077	1 AT10 HUMAN	Q9h324 homo sapien
3	1207	23.1	860	1 AT56 HUMAN	Q9ukp5 homo sapien
4	1205.5	23.0	1593	1 AT12 HUMAN	P58397 homo sapien
5	1189	22.7	997	1 AT57 HUMAN	Q9ukp4 homo sapien
6	1086.5	20.8	1306	1 AT20 MOUSE	P59511 mus musculus
7	1061.5	20.3	1629	1 AT59 HUMAN	Q9p2n4 homo sapien
8	1027.5	19.6	1911	1 AT20 HUMAN	P59510 homo sapien
9	1022.5	19.5	968	1 AT51 MOUSE	P97857 mus musculus
10	1021.5	19.5	967	1 AT51 RAT	Q9wq1 rattus norv
11	1018.5	19.5	1207	1 AT19 HUMAN	Q8te59 homo sapien
12	998	19.1	1205	1 AT53 HUMAN	O15072 homo sapien
13	989	18.9	1210	1 AT19 MOUSE	P59509 mus musculus
14	985	18.8	1095	1 AT17 HUMAN	Q8te56 homo sapien
15	984.5	18.8	367	1 AT51 HUMAN	Q9un18 homo sapien
16	948	18.1	1205	1 AT52 BOVIN	P79331 b adants-2
17	948	18.1	1211	1 AT52 HUMAN	O95450 h adants-2
18	929.5	17.8	950	1 AT15 HUMAN	Q8te58 homo sapien
19	887	16.9	1223	1 AT14 HUMAN	Q8wx58 homo sapien
20	886.5	16.9	930	1 AT55 MOUSE	Q9r001 mus musculus
21	878	16.8	930	1 AT55 HUMAN	Q9una0 homo sapien
22	870.5	16.6	630	1 AT54 RAT	Q9espf rattus norv
23	856.5	16.4	837	1 AT54 HUMAN	P57173 homo sapien
24	826.5	15.8	905	1 AT58 MOUSE	P71110 mus musculus
25	820.5	15.7	890	1 AT58 HUMAN	Q9up79 homo sapien
26	698	13.3	562	1 AT15 MOUSE	P59384 mus musculus
27	405	7.7	207	1 AT55 BOVIN	Q9tt92 bos taurus
28	340.5	6.5	245	1 AT54 BOVIN	Q9tt93 bos taurus
29	334.5	6.4	920	1 AD19 MOUSE	O35674 mus musculus
30	326	6.2	450	1 AT10 MOUSE	P58459 mus musculus
31	316	6.0	956	1 AD19 HUMAN	Q9h013 homo sapien
32	308.5	5.9	824	1 AD08 HUMAN	P78325 homo sapien
33	305	5.8	826	1 AD08 MOUSE	Q05910 mus musculus

ALIGNMENTS

RESULT 1

ID	AT18 HUMAN	STANDARD;	PRT;	1081 AA.
AC	Q8TE60;			
DT	15-SEP-2003 (Rel. 42, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DE	ADAMTS-18 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 18) (ADAM-TS 18) (ADAM-TS18).			
GN	ADAMTS18.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21855482; PubMed=11867212;			
RA	Cal S., Ohaya A.J., Llanazares M., Garabaya C., Quesada V., Lopez-Otin C.;			
RT	"Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains.";			
RL	Gene 283:49-62(2002)			
CC	!- COPACTOR: Binds 1 zinc ion per subunit (By similarity).			
CC	!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).			
CC	!- TISSUE SPECIFICITY: Expressed in fetal lung, liver, and kidney and in adult brain, prostate, submaxillary gland, and endothelium.			
CC	!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).			
CC	!- SIMILARITY: Belongs to peptidase family M12B.			
CC	!- SIMILARITY: Contains 1 disintegrin-like domain.			
CC	!- SIMILARITY: Contains 4 TSP type-1 domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
DR	EMBL; AJ311903; CAC83612.1; -			
DR	Genew; HGNC:17110; ADAMTS18.			
DR	MIM; 607512; -			
DR	InterPro; IPR002870; Pep_M12B_propep.			
DR	InterPro; IPR001590; Reprolysin.			
DR	InterPro; IPR000884; TSP1.			
DR	Pfam; PF01562; Pep_M12B_propep; 1.			
DR	Pfam; PF01421; Reprolysin; 1.			
DR	Pfam; PF00090; tsp1; 4.			
DR	SMART; SM00209; TSP1; 4.			
DR	PROSITE; PS0215; ADAM_MEPRO; 1.			
DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.			
DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.			
DR	PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.			

34	281	5-4	857	1	AD22_MOUSE	Q9rlv6 mus musculus
35	269.5	5.1	616	1	ECAR_ECHCA	Q90495 echis carin
36	260.5	5.0	819	1	AD09_HUMAN	Q13443 homo sapien
37	260	5.0	903	1	AD12_MOUSE	Q61824 mus musculus
38	254	4.9	906	1	AD22_HUMAN	Q9p0k1 homo sapien
39	253.5	4.8	697	1	AD26_MOUSE	Q9rl58 mus musculus
40	252	4.8	729	1	AD21_MOUSE	Q9rl76 mus musculus
41	245.5	4.7	813	1	AD33_HUMAN	Q9bz11 homo sapien
42	243.5	4.7	814	1	AD15_HUMAN	Q13444 homo sapien
43	238	4.5	1077	1	SM5A_MOUSE	Q62217 mus musculus
44	237.5	4.5	775	1	AD28_HUMAN	Q9ukq2 homo sapien
45	237.5	4.5	909	1	AD12_HUMAN	O43184 homo sapien


```

DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS00142; ZINC PROTEASE; FALSE NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 47
FT PROPEP 48 284
FT CHAIN 285 1081
FT DOMAIN 285 497
FT DOMAIN 285 497
FT DOMAIN 498 577
FT DOMAIN 588 643
FT DOMAIN 644 749
FT DOMAIN 750 876
FT DOMAIN 877 931
FT DOMAIN 933 991
FT DOMAIN 993 1048
FT SITE 254 254
FT METAL 435 435
FT ACT_SITE 436 436
FT METAL 439 439
FT METAL 445 445
FT CARBOHYD 151 151
FT CARBOHYD 190 190
FT CARBOHYD 313 313
FT CARBOHYD 744 744
FT CARBOHYD 837 837
FT CARBOHYD 908 908
FT CARBOHYD 1081 AA; 1.9655 MW; 0438BF6456764616 CRC64;
SQ SEQUENCE 1081 AA; 1.9655 MW; 0438BF6456764616 CRC64;

Query Match 46.0%; Score 2408; DB 1; Length 1081;
Best Local Similarity 51.0%; Pred. No. 1.7e-167;
Matches 463; Conservative 118; Mismatches 197; Indels 130; Gaps 18;

QY 85 ASSSTRGAGLDGKGRDNDAGNHRSSQNTGTNTQTLHLVLTQYDLSAYEVDHGRGYS 144
DB 47 ASDSSGSGLN-----DDYFVFPVEVDGAGSYIS 77

QY 145 HEIMHQRARRAVAVESESLHLKGRPHDMLDRTSSLSVAFGTVQLGTGKSV 204
DB 78 HDLLHNGKKRS-AQNARSSLHYRFSAGQELHLELK-PSALLSHFIVQLGKGASSET 135

QY 205 QTLPEDECFVQSLRSHRNPSHGKFCGEGSTRTLKLNQ-----KCPDSDVFRAQC 260
DB 136 QK-PEVQCFYQGFIRND-SSSSAVSTCAGLSGLIRTRKNEFLISPLFQ-----LLA 186

QY 261 AEHNSRRFRGRHYKWKPYTOVEADLCKLYCIAEGDFPFFLSNKKVQD-----TPCEDS 315
DB 187 QEHNYSSPAG-HHPVLVYKRTAEKIQRYVFG-----SGRNYFGYSFHSHPHASQS 238

QY 316 R-----NVCIDGICELSVSTSAHMPQPKEDLFLPDEYKSCURHKRSLRS 363
DB 239 RETEYHRRRLQKHPC-----GRRKKVAPKPTEDTYLRFDEYSGSRPRRSAGKS 289

QY 364 HNEELNVETLVVDKQKQVGHENITTYVLTILNMVYSALEFKOG----- 408
DB 290 QKG--LNVETLVADKRWKXHGKGNVTYLTWNK-VSGLFKQGTIGSDINNVVLSIL 346

QY 409 -----LMGDKGTRHDHAILLTGLIDCSWKNKPCDTLGF 441
DB 347 LEQEPGGLLINHHADQSLNSFCWQSALIGKNGRKHDAILLTGFDICSWKNKPCDTLGF 406

QY 442 APISGCKYKRSCTINEDTGLGTAFTIAHSGHNFPMTHDGEKGNCKKSEGNIMSPITAG 501
DB 407 APTSGMCKYKRSCTINEDTGLGTAFTIAHSGHNFPMTHDGEKGNCKKSEGNIMSPITAG 466

QY 502 RGVFESWPCSRQYLLHKLFLSTAQAICLADOPKFKVYKPEKLPCELYDANTCKWQFGE 561
DB 467 NNGVFSWSSCSQYLUKFLSTPQAGCLVDEFAQAGQYKPKLPQIYDADTQCKWQGA 526

QY 562 KAKLQMLDKDICKALMCHRIKRCCKETKFMFAEGTICGHDWRCGQCVKYGDEGPKP 621
DB 527 KAKLSLGFVQDKICKSLNCHRVHRCCKETKFMFAEGTICGHDWRCGQCVKYGDEGPKP 586

QY 622 THGHSWDSWSSPCSRCTGGGVSHRSRLCTNPKPSHGKFCFEGSTRTLKLNCKKCRPDS 681

```

```

DB 587 IHQWQAWKSNSECSITCGGVKFKQHRHCNNPAPQYGGIFCGSSRIYQLCNINPCNS 646
QY 682 VDFRAAQCAEHSRRFRGRHYKWKPYTOVEADLCKLYCIAEGDFPFFLSNKKVQDGTTC 741
DB 647 LDFRAQCAEYNSKPFGRFWYQWKPYTKVEEDRCKLYCKAENFEFFFAFGSKVKGDTTC 706
QY 742 SEDSRVNCIDGICERVCGDNVLGSDADEVCGVCGNNSACTIHRGLYTKHHHTNQYHM 801
DB 707 SPKNVNCIDGVCELGVGDHGLSKAVSDACGVCKGDNSTCKFYKGLYLQHKANBYYPV 766
QY 802 VTIPSGARSIRIYEMNVSTSYISVRNARLYLNGHWTVDMPGRYKFSGTFDYRRSYNE 861
DB 767 VLIIPAGARSIEIOELQVSSSYLAIVRSLSQXYLTGWSIDWDFGEPFAGTTFEYQSFNR 826
QY 862 PENLIATGPTNETLIVELLFQGRNPGVAMEYSPRL--GTEKQPPA--QBSYTWAIIVRSE 917
DB 827 PERLYAPGPTNETLIVELLFQGRNPGVAMEYSPRL--GTEKQPPA--QBSYTWAIIVRSE 917
QY 918 CSVSCGGG 925
DB 884 CSVSCGGG 891

RESULT 2
AT10 HUMAN STANDARD; PRT; 1077 AA.
ID AT10 HUMAN STANDARD; PRT; 1077 AA.
AC Q9H324;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-10 precursor (SC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 10) (ADAM-TS 10) (ADAM-TS10) (Fragment).
DE ADAMTS10.
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Apte S.S.;
RT "ADAM-TS10: a novel member of the ADAM-TS family containing multiple
RT thrombospondin type 1 repeats.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -! COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -! SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -! DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (By
CC similarity).
CC -! SIMILARITY: Belongs to peptidase family M12B.
CC -! SIMILARITY: Contains 1 disintegrin-like domain.
CC -! SIMILARITY: Contains 1 PLAC domain.
CC -! SIMILARITY: Contains 5 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF163762; AAC35563.1; -.
CC MEROPS; M12.235; -.
CC Genew; HGNC:13201; ADAMTS10.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR002870; Pep_M12B_propep.
CC InterPro; IPR001590; Repolysin.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR006025; Zn_WTpeptdase.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Repolysin; 1.

```


DR Pfam; PF00030; tss_1; 5.
 DR SMART; SM00209; TSP1; 5.
 DR PROSITE; PS00215; ADAM_MEPRO; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE; PS00092; TSP1; 5.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR HydroLase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix.
 FT NM TER 1
 FT PROPEP <1 207
 FT CHAIN 208 1077
 FT DOMAIN 208 433
 FT DOMAIN 434 520
 FT DOMAIN 578 679
 FT DOMAIN 680 802
 FT DOMAIN 521 576
 FT DOMAIN 799 857
 FT DOMAIN 858 919
 FT DOMAIN 921 975
 FT DOMAIN 977 1032
 FT DOMAIN 1043 1074
 FT METAL 366 366
 FT ACT SITE 367 367
 FT METAL 370 370
 FT METAL 376 376
 FT METAL 64 64
 FT CARBOHYD 196 196
 FT CARBOHYD 297 297
 FT CARBOHYD 714 714
 FT CARBOHYD 769 769
 FT CARBOHYD 866 866
 SQ SEQUENCE 1077 AA; 118072 MW; 3914DE18DCBBF587 CRC64;

Query Match 23.48; Score 1223; DB 1; Length 1077;
 Best Local Similarity 31.28; Pred. No. 3.4e-81;
 Matches 283; Conservative 135; Mismatches 302; Indels 186; Gaps 22;

QY 117 TENQTLVLTQYDLVSAYVDHGDYVSHMEHRRRAVAVSEVSHLRKGRPHDF 176
 DB 2 SQDEFLSSLEYETAPTRVDHNGALLAFSPPPRRQRTGATASRLFYKVASPTHF 61
 QY 177 HMDLRTSSSLVAGFIVQTLGKTGTSVQTLPEDEFCFGSLRSRHSN----- 225
 DB 62 LNLNTRSSRLLAGHSVEYWTREGLAWQRAAPH--CLYAGHLQGOASSSHVAISTCGSL 119
 QY 226 -----PSHGGRFCGSTRTLKLCNSQKCPDSS--VDFRAAQ----- 259
 DB 120 HGLIVADBEELIPLHGGP--KGS-----RSPESGPHVVVKRSLRHPHLDTA 167
 QY 260 CAEHSRRFRGRHYKWKPYQVEADLCKLYCIAEGDFPFLSKNVKDGTPCSEDSRNV 319
 DB 168 CGVRDEKPKWGRPMWLTLLKPPAR-----PLGNETERGPP----- 203
 QY 320 IDGICELSVSTSAHMPQPKEDLFIPLDEYKSLRSLRSHRNEELNVETLVVDK 379
 DB 204 --GL-----KRSVR-----EYVETLVVDK 223
 QY 380 KQMNHGHENITTVYILNMVSALFKDGLMG----- 411
 DB 224 MWAYHGRDVEQYVLAMINIVAKLFQDSLSGTNVILVTRLILTQOFTLEITHHAGK 283
 QY 412 -----KDG-TRHDHAILTGLDICSWMKNEPCDTLGFAPISM 447
 DB 284 SLDSFCKWQKIVHSHGHGNAIPENGVAHNDTAVLIRYDILCYKPKCTGLAPVGM 343
 QY 448 CSKYRSCNTNBDTGLGTAFTIABSGHFMHDEGNMC---KKSEGNIMSTPLAGRN 504
 DB 344 CERERSCSVNEDIGLPQAFTIAEIGHTFGWNHGVGNSCGARGQDPKALMAAHITMKTN 403
 QY 505 VFSVSPCSRQLHKLFLSTAQACLADQPKVKEKYKPKLPCELYDANTOCKQFGEKAK 564
 DB 404 PFWSSNCNRDYITSFLDSGLGLCLNNRP-PKQDFVYPTVAPGQAYDADEQCQFHGVKSR 462

QY 565 LCMDFKFKDICKALWCHRIKCKCETKFMFAAESTIC-GHDM----WCRGGCQKYKGDGPK 620
 DB 463 QCKYK-----EVCSELWCLSKSNRCITNSIPAAEGLTCQHTIDKGCYKRVCFVFSR-PE 518
 QY 621 PTHGHSWSSWSPCSTRTCGGVSHSRSLCTNPKPSHGKFCGSTRTLKLCNSQKCPRD 680
 DB 519 GVDGAWGPWTPNGCSTRTCGGVSSSRHCDSPRTPIGGKYCLGERRRHSCNTDPCPG 578
 QY 681 SVDFAAQAACHNSRRFRGRHYKWKPYQVEODDLCKLYCIAEGDFPFLSKNVKDGTP 740
 DB 579 SQDFREVQCESEFSDIPRGRKFKWKY-RGGGVKACSLTSLAEGFNFYTERAAVVDGTP 637
 QY 741 CSEDSRNVICDGCERVCNVLGSDAIVEDVCGVGNNSACTIHRGLTYKHHTNQYH 800
 DB 638 CRPDTVDICVSGECKHVGCDRLVGLSDUREDKRCVCGDGSACETIIEGVFSPAGCYED 697
 QY 801 MVTIFSGARSIRIVEMNVSTSYISVRNALRRYLYNGHWTVDMPGRYKFGSTTFDYRRSYN 860
 DB 698 VVMIPKGSVHIFIQDLNLSLHALQDQBSLLLEGLPGTPQPHRLPLAGTTFLQLRQGP 757
 QY 861 EPENLATGPTNETLIVELLFOGNSGVAVEYSMBRLGTGTEKQPPAQPSYTWIVR-SECS 919
 DB 758 QVQSGLEALGPNASLIWVLTARTELPALRYFRNAP-IAROSLFP-----YSNVYAPWKCS 812
 QY 920 VSCGGG 925
 DB 813 AQCAGG 818

RESULT 3
 AT56 HUMAN
 ID AT56 HUMAN STANDARD; PRT; 860 AA.
 AC Q9UKF5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ADAMTS-6 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 6) (ADAM-TS 6) (ADAM-TS6).
 GN ADAMTS6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9395124; PubMed=10464288;
 RA Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;
 RA "ADAM-TS6, ADAM-TS6, and ADAM-TS7, novel members of a new family of
 zinc metalloproteases";
 RL J. Biol. Chem. 274:25555-25563 (1999).
 CC -!- COPACITOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN PLACENTA AND BARELY
 CC DETECTABLE IN A NUMBER OF OTHER TISSUES.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF140674; AAD56357.1; -.

DR HSP: P34179; LIAG.
 DR KEROPS; M12.230; --
 DR Genew; HGNC:222; ADAMTS6.
 DR MIM; 605008; --
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR002870; Pep M12B propop.
 DR InterPro; IPR001590; Repolysin.
 DR InterPro; IPR000884; TSPI.
 DR InterPro; IPR006025; Zn_Mtpetpdee.
 DR Pfam; PF01562; Pep M12B propop; 1.
 DR Pfam; PF00090; Repolysin; 1.
 DR SMART; SM00209; TSPI; 1.
 DR PROSITE; PS00215; ADAM_MPRO; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00092; TSPI; 1.
 DR PROSITE; PS00427; DISINTEGRIN 1; FALSE NEG.
 DR PROSITE; PS0214; DISINTEGRIN 2; FALSE NEG.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix.
 FT SIGNAL 1 21
 FT PROPEP 22 244
 FT CHAIN 245 860
 FT DOMAIN 245 452
 FT DOMAIN 453 509
 FT DOMAIN 510 565
 FT DOMAIN 567 668
 FT DOMAIN 669 795
 FT DOMAIN 796 852
 FT DOMAIN 68 71
 FT DOMAIN 662 665
 FT METAL 403 403
 FT ACT SITE 404 404
 FT METAL 407 407
 FT METAL 413 413
 FT CARBOHYD 99 99
 FT CARBOHYD 172 172
 FT CARBOHYD 222 222
 FT CARBOHYD 234 234
 FT CARBOHYD 676 676
 FT CARBOHYD 843 843
 SQ SEQUENCE 860 AA; 97098 MW; E57213015DEC82C5 CRC64;
 Query Match 23.1%; Score 1207; DB 1; Length 860;
 Best Local Similarity 31.7%; Pred. No. 3.8e-80;
 Matches 286; Conservative 131; Mismatches 291; Indels 194; Gaps 25;
 QY 108 HRSQQTNTGTENQTLHLVLTQYDLVSAYEVDHREGDYVSHEIMH--HQRERRAV----AVSE 161
 DB 22 HSDRLSYSSQEEFLYLEHYQLTPIRVQDQGAFLSFTVNDKHSRRRSNDPDPQQA 81
 QY 162 VESLHLRLKGRPHDMOLRTSSLSVAPQIVQTLGKTGTSVQTLPE-----DQCFY 215
 DB 82 VSKLFFKLKSAVKGKPHNLTLNTDFVSKHFTVEYWGKG-----PQWKHDFLNCXY 133
 QY 216 QGSLRSHRNSPSHGKFCGEGSTRTLKLNQSKCPDRSDVF-----RAACQAEHNSRRFRG 270
 DB 134 TGYLDQDRSTTKVALSNVGLHCVI-----ATEDEEYFIEPLKNTTDSKHSYENGH 186
 QY 271 RHYKWKPYQVQVADLC-KLYCIAEGDFDFSSLSNKKVKGTP-CSDSRNVICDICELSV 328
 DB 187 PHVIYKSAQQRHLYDHSCH-----GVSDTRSGKFWNLNDTSVY----SLPI 233
 QY 329 VSTSAHMQPPKEDLFIPLDEYKSLRHKRSLRSHRNEELNVVLVVDKKNQNHGHE 388
 DB 234 NNTHIHH-----RQKRSV-----SIERFVETLVVADRMVGYGHRK 269
 QY 389 NITTVVLILNMYSLPFQDLGM----- 411
 DB 270 DIELVILSVMIKVALKYLSDSILGNVYVITVARLVLITBEDQPNLEINHAKSLDSFCKWQ 329
 QY 412 -----KDG-----TRHDHALLTGLDICSWKNEPCCDTLGFAPISGWCYSKRSCTI 456
 DB 11
 Db 330 KSILSHQSDGNTIPENGIAHNDNAVLITRYDICTYKNKPCGTLGLASVAGMCEPERSCSI 389
 QY 457 NEDTGLGLAFTIAHSGHNFQMIHDGEGNMCKSEGNIMSPTLAGRNQVFSNPSQRYQL 516
 Db 390 NEDIGLSAFTIAHEIVHFGNHDGICNSC-----GR----- 422
 QY 517 HKFLSTAQAICLADQPKPVKEYKPEKULGEBLYDANTOCKWQFGKAKLCLM-----DFK 571
 Db 423 -----KVMKQ-----QNYGSSHYCEYQ--SFFLVCLQSRXHHQLF 455
 QY 572 KDICKALACHRIGRKCETKEMPAABGTICG-----HDMWCRGCGQCKYKDEGPKPTHGWS 627
 Db 456 REVCKELMCLSKSNKRCVTNSIPAAEGTLCQTGNIKMGWCYQDCDVPFG--TPQSDGGWG 514
 QY 628 DWSWSPCSRTCGGGVSHRSRLCTNPKPSHGKFCGEGSTRTLKLNQSKCPDRSDVDFRAA 687
 Db 515 FWSLWGECSRTCGGVSSSLRHCDSPAPSGGKCYCLGERKYRSCNTDPCPLGSRDFREK 574
 QY 688 QCAEHNSRRPRGRHYKWKPYQVEDQDLCKLYCIAEGDFDFSSLSNKKVKGDTGPTCEDSRN 747
 Db 575 QCADFDNNPFRGKYNNKPYTGGGVKPK-CALNCLAEAGYNYTERAPAVIDGTQCNADSLD 633
 QY 748 VCIDQICERVGCNVLGSDAIVEDVCGVCNNGNSACTIHRGLYTKHHTNQYHYHMTVPSG 807
 Db 634 ICINGECKHVGDNLGSDAREDCRVCGGGGSGTCDAGEFNDLSLPRGGYWEVVQIPRG 693
 QY 808 ARSRIYEMNVSTSYISVRNALRYLYLNGHWTVDPGRYKPSGTTDFYRRSYNBPENLIA 867
 Db 694 SVHIEVREAVAMSKYIALKSEGDYIYINGAWTIDWPKRFDVAGTAFHYKRPDTBPESLEA 753
 QY 868 TGPNTETLIVELLFQGRNPGVAWEYSMP--RLGTETKQPPAPQSVYTAIVR--SECSVSCGG 924
 Db 754 LGPTSENILVNLQEQNLGIRYFNFPITRTGSDN---EVGFTWNHQPWSECSACAG 810
 QY 925 GR 926
 Db 811 GK 812
 RESULT 4
 AT12 HUMAN
 ID AT12 HUMAN STANDARD; PRT; 1593 AA.
 AC P58397;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12).
 GN ADAMTS12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=11279086;
 RA Cal S. Argueelles J.M., Fernandez P.L., Lopez-Otin C.;
 ET "Identification, characterization, and intracellular processing of ADAM-TS12, a novel human disintegrin with a complex structural organization involving multiple thrombospondin-1 repeats.";
 RT J. Biol. Chem. 276:17932-17940 (2001).
 RL -I- COPACTOR; BINDS 1 ZINC ION (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).
 CC -I- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely expressed in gastric carcinomas and in cancer cells of diverse origin.
 CC -I- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (B; SIMILARITY).
 CC -I- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
 CC -I- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ADAMTS-7 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 7) (ADAM-TS 7) (ADAM-TS7).
 GN ADAMTS7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93395124; PubMed=10464288;
 RA Hurskainen T.L., Hirohata S., Seidin M.F., Apté S.S.;
 RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
 RT zinc metalloproteinases";
 RL J. Biol. Chem. 274:25555-25563(1999).
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,
 CC LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 1 PLAC domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF140675; AAD56358.1; -;
 DR HSPP; P15167; 1ATL.
 DR MEROPS: M12.231; -;
 DR Genew: HGNC:223; ADAMTS7.
 DR MIM: 605009; -;
 DR InterPro: IPR001762; Disintegrin.
 DR InterPro: IPR002870; Pep_M12B_propep.
 DR InterPro: IPR001590; Reprolysin.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR006025; Zn_MTPeptidse.
 DR Pfam: PF01562; Pep_M12B_propep; 1.
 DR Pfam: PF01421; Reprolysin; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR SMART: SM00209; TSP1; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR PROSITE: PS00215; ADAM_MPRO; 1.
 DR PROSITE: PS00545; CYSTEINE_SWITCH; FALSE_NEG.
 DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE: PS00214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE: PS00092; TSP1; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix.
 FT SIGNAL 1 27
 FT PROPEP 28 232 POTENTIAL
 FT CHAIN 233 997 BY SIMILARITY.
 FT DOMAIN 233 461 ADAMTS-7.
 FT DOMAIN 462 537 METALLOPROTEASE.
 FT DOMAIN 538 593 DISINTEGRIN-LIKE.
 FT DOMAIN 595 697 TSP TYPE-1 1.
 FT DOMAIN 698 914 CYS-RICH.
 FT DOMAIN 915 990 SPACER.
 FT DOMAIN 944 974 TSP TYPE-1 2.
 FT DOMAIN 204 204 PLAC.
 FT SITE 204 204 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 389 389 BY SIMILARITY.

FT METAL	392	392	ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL	398	398	ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD	94	94	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	693	693	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	778	778	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE	997 AA; 109694 MW; 6587044ED02FC104 CRC64;		

Query Match 22.7%; Score 1189; DB 1; Length 997;
 Best Local Similarity 30.5%; Pred. No. 9.3e-79;
 Matches 294; Conservative 125; Mismatches 305; Indels 240; Gaps 26;

QY	129	DLVSAVEVDHGRGDYVSHHMHQRRAVAVSEVSLHLRLKGRPHDFHMDLRTSSSLVA	188
DB	44	DIVHVRVDSAGSFLSYELWPRALRKROVSRVRRDAPAFYELQVGRGELRNFNLTAHQHLLA	103
QY	189	PGFIVOTLGTGKTSKVOTLPDPDFCFYQGLSRSHRNPSHGG- - -KFCGSGTRTLKLCN	244
DB	104	PGFVSTRRRGGLGRAHRAHTPACHLLGEV- - -QDPELEGGALAIACDGLKGVQLSN	160
QY	245	SQK- - -PRDSVDFRAQCAEH- - -NSRRFRGRHYKWKPYT- - -QVEADLCKLYCIAE	293
DB	161	EDYFIEFLDSAPARPGHAQPHVYKROAERLAQRGDSAPSTCGVQVYPEL- - -	212
QY	294	GFDFFFLSNKVKDGTGPTCEDSRNVICIDGICELSVVSTSAHWPQPKDELFLPDEYKSC	353
DB	213	- - - - -ESSR- - - - -ERWEQR	222
QY	354	LRHKSLLRS-HR- - -NEELNVETLVVVKQKMQHGHENITTYLTINMVSALFKGL	409
DB	223	QWRRLRLRLQRSVKEKVVETLVADAKMVEYHGQPVESYVLTINMVAFLHDPDS	282
QY	410	MKG- - - - -DGT- - - - -RHDHAILLT	424
DB	283	IGNPHITVRLVLEDEEDLKIITHAUNTUKSCFKWOKSNMKGDAHLHHDHAILLT	342
QY	425	GLDICSWKNEPCTDLGFAPISQCMCKYRSCTINEDTGLGLAFTIAHESGHNFQIMHDEG	484
DB	343	RKDLCAAMNRPCTGLSHVAGMCPHRSCSINEDTGLPLAFTVAHELGHSGFIQHDGSG	402
QY	485	NWCKK- - -SEGNIMSPFLAGNGVFWNSPCSROYLHKFLSTAOAI CLADOP- - -KPVKYKYP	541
DB	403	NDCPVGKRPFTMSPOLLYDAAPLMTWSCRSQYITRFLDRGWLGLDDPPAKDIID- - -FP	460
QY	542	EKLPGELYDANTQCKWQFCEKAKLMDPKDICALWCHRIKRCCTKFMFAAEGTICG	601
DB	461	SVPPGVLYDVSHQCRLOYGAYSAC- - -EDMNVCHTLAC- - -SVGTTCHSKLDAVDTGRCG	517
QY	602	HDWRCRGQCVKYGDGPKPTHGHWSDWSSWSPSCSTCGGVSHRSRLCTNPKPSHGKRF	661
DB	518	ENKWCILSGECVPVGF- - -PEAVDGGWGSAMWSICSRSCGMGVQSAERQCTQPTPKYKRY	576
QY	662	CEGSTFTLKLNSOKCPRDSVDFRAQCAEHNSRRFRGRHYKWKPYTQVEDODLCKLYCI	721
DB	577	CVGERKRFRLCNLQACPAGRPSPFRHVQCDSHFDFAMLYKQLHTWVPV- - -VNDVNPCELHCR	634
QY	722	ARGFOFFFLSNKVKDGTGPTCSE- - -DSRNVICIDGICERVGCDNVLGSDADEVCGVCNGNN	779
DB	635	PANEYFAKLRDAVDGTFQVRSRDLCTINGICNVGCDPEIDSGAMEDCGVCHGNG	694
QY	780	SACTHRLGLYTKHHTNQYHNVTTIPSGARSTRIRIYEMNVSTSYISVRNA- - -LRRYLYNGHW	838
DB	695	STCHTVSGTFFEEAEGLG- - -YVDVGLIPAGAREIRIQEVAEAAANFLALRSDEPKYFLNGW	753
QY	839	TVDWPCRYKFSQTFDTRRSYNEPENLIATGPTNE- - -	873
DB	754	TIQWNGDIQVAGTITFYARRGNW- - -ENLTSPGTKEPVMIQVPASRGPGSGRGVPRPST	812
QY	874	- - - - -	873
DB	813	LHGRSRPGVSGSVTEPCSEPGPPAAASTSVPSLKWPNLVAHVHGRGWGAPILGLGW	872
QY	874	- - - - -TLIVELLFQGRNPGVAWEYSMPRL- - -GTEKQPPAQPSYTWAI- - -VRSECSVS	921

Db 873 RRHLVLMGRLPTQLLFOESNPGVHYEYTHREAGGHDEVP--PVFSWHYGPWTKCTV 930

QY 922 CGGG 925

Db 931 CGRG 934

RESULT 6

ID AT20 MOUSE STANDARD; PRT; 1906 AA.

AC P59311; STANDARD; PRT; 1906 AA.

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-TS20).

GN ADAMTS20.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]_TaxID=10090;

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal brain.

RA Llanazares M., Cal S., Lopez-Orin C.;

RT "Identification and characterization of ADAMTS-20 defines a novel subfamily of metalloproteinases-disintegrins with multiple thrombospondin-1 repeats and a unique GON-domain."

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: May play a role in tissue-remodeling process occurring in both normal and pathological conditions.

CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -!- TISSUE SPECIFICITY: Expressed at low level in testis and brain.

CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -!- SIMILARITY: Belongs to peptidase family M12B.

CC -!- SIMILARITY: Contains 1 disintegrin-like domain.

CC -!- SIMILARITY: Contains 1 GON domain.

CC -!- SIMILARITY: Contains 15 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL; AJ512753; CAD54808.3; -.

DR SMART; SM00209; TSP1; 14.

DR PROSITE; PS00215; ADAM_MBP1; 1.

DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.

DR PROSITE; PS00214; DISINTEGRIN_2; FALSE NEG.

DR PROSITE; PS00092; TSP1; 13.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;

KW Repeat; Extracellular matrix.

FT SIGNAL 1 26 POTENTIAL.

FT PROPEP 27 249 BY SIMILARITY.

FT CHAIN 250 1906 ADAMTS-20.

FT DOMAIN 250 464 METALLOPROTEASE.

FT DOMAIN 465 552 DISINTEGRIN-LIKE.

FT DOMAIN 553 608 TSP TYPE-1 1.

FT DOMAIN 609 720 CYS-RICH.

FT DOMAIN 721 842 SPACER.

FT DOMAIN 843 901 TSP TYPE-1 2.

FT DOMAIN 906 962 TSP TYPE-1 3.

FT DOMAIN 962 1015 TSP TYPE-1 4.

FT DOMAIN 1017 1074 TSP TYPE-1 5.

FT DOMAIN 1075 1131 TSP TYPE-1 6.

FT DOMAIN 1148 1202 TSP TYPE-1 7.

FT DOMAIN 1203 1260 TSP TYPE-1 8.

FT DOMAIN 1300 1351 TSP TYPE-1 9.

FT DOMAIN 1354 1411 TSP TYPE-1 10.

FT DOMAIN 1412 1465 TSP TYPE-1 11.

FT DOMAIN 1468 1526 TSP TYPE-1 12.

FT DOMAIN 1527 1584 TSP TYPE-1 13.

FT DOMAIN 1585 1648 TSP TYPE-1 14.

FT DOMAIN 1650 1706 TSP TYPE-1 15.

FT DOMAIN 1707 1906 GON.

FT METAL 399 ZINC (CATALYTIC) (BY SIMILARITY).

FT ACT SITE 400 BY SIMILARITY.

FT METAL 403 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 409 ZINC (CATALYTIC) (BY SIMILARITY).

FT CARBOHYD 92 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 221 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 714 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 798 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 805 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1057 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1562 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1719 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1759 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1777 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;

Query Match 20.8%; Score 1086.5; DB 1; Length 1906;

Best Local Similarity 30.6%; Pred. No. 6.1e-71;

Matches 276; Conservative 146; Mismatches 302; Indels 177; Gaps 37;

QY 125 LTQYDLVSAYEVDHGRDGVSHHMHQRRRAVAVSEVESL--HLRLKGRPHDFHMDLRT 182

Db 37 LASVEVWPTRVNEPGD-VFQNRHFSKRRSSGVPEPPFRTHYRISAYGOLFQLNLSA 95

QY 193 SSSLVAPGFIVQTICK-----TGTSKVQTLPPDEFQCYGSLSRHNSHSGKFP--CEGS 236

Db 96 DAFLAAGYTEVHLGTVPVPGGRSTES-PDLRHCFYRGQVNARED---HTAVFSLCGSL 151

QY 237 TRTLKLCNSQKCPDRSDVDFRA---AQCAEHSRRRGRHRYKWKVPYQVADLCKLYCAE 293

Db 152 MGTFK-ANDGEVLEPV-LRADGSAHDDDN-----KPHLIYRQELKR----- 192

QY 294 GDFPFPSLSNKKVDGTPC-----SEDSRNVICIDGICELSVSVSTSAHMPQPK 340

Db 193 ---NSFARSHK-----PCEVSENOEMKALPQSSRNTTGDVDIE-----E 230

QY 341 EDLFLPDEYKSLRHKESLLSRHNEELN---VETLVVVDKKMKQKHENITYVLT 396

Db 231 EAVFRLGE-----RSQLHSRNKRFSLSPRYVEVMVTADAKMNVHHG-QNLQHYVLT 281

QY 397 ILNMYSAFLFKDGLMG-----KDG----- 414

Db 282 LMSIVAALYKDDSSIGNLINIVIKLVTHSEQGPVIFSNAAITLRFNCLWQSQNVPPD 341

QY 415 ---TRHDHAILLTGLDICSNNNEPCDITLGPAPIGOMSKYKSCSTINEDTGLAFTIAHE 471

Db 342 AHPSHHDTAVLITREDICGAK-EKCDITLGLAELGTLCDPSRSCSISEENGLSAAFTIAHE 400

QY 472 SGHNFGMTHDGEGNMCK---KSEGNINSPITLAGNNGVFSWSPCSRQVLHFLSTAQAI 527

Db 401 LGHVFNVPD--DSPKCKEAGIKHQHVMAPTLNTYHTSPWTWSACSQKHITFELDTGHGEC 459

QY 528 LADQPKPKYKYPEKLPGELYDANTQCKWQFGKAKLCLMDFKCKICKALWC---HRIG 584

Db 460 LLDKPNG-RTYDLSPLQPLGVSVDGNRQCELMFPGSQVC--PYLKH-CRRLWCTSAEGVH 515

QY 585 KCBTKFMPAAEGTICQHDMMCRGCGQCVKYQDEGPKPTHGHWSDWSNSPSCRTRCGGVS 644

Db 516 KGCRTQHMPLADGTSCGPGMHCHRGCLVT-RDMETRPVDGEGWPGWPGVSSCRTCGGG 574

QY 645 HRSRLCTNPKPSHGKFCGEGSTRTLKLCNSQKCPDRSDVDFRAAOCAEHSRFRGRHY-- 702

Db 575 STARLCORPEPRNGRGYCVGRMKFRKSCNTDSCPKGKRDREKQCSD-----FDGKHFDI 629


```

QY 703 -----KWKP-YTQVEDODLCKLYCIAEGDFPFSLNKNVKGDTGPPCSEDSRNVICDGI 754
DB 630 NGLPENVRLPKYSYGIKAVKDRCKLYCRVAGTTFYQLKDRVADGTFCTGTNDICVQGLC 689
QY 755 ERVGDNLVGSDAVEDVCGVNGNNSACTIHRGLYTKHHHTNQYVHMVTIPSGARSIRY 814
DB 690 ROAGCDHVLNSAKKDKCKCGVCGDSSQTLAGVNSAHY--GYNVVKIPAGATNIEL 747
QY 815 ENNV-----TSYISVRNALRRYLYNGHWTVDWPGR-YKFSGTTFDYRRSYNEPENLIAT 868
DB 748 QHSYSGRPDDNYLALSDTQGNFLNGNFVSMKKEINIQAFAVEYSGSNNSIERINST 807
QY 869 GPTNETLIVELLFQGR--NPGVAMEYSMP---RLGTEKOPPAQPSYTWAIVRSECSVSCG 923
DB 808 DRLEAEALQVLQVLCVNLNPDVRYSFNPIEERSNLFSWDYFYP---W-----QDCTRMQ 860
QY 924 G 924
DB 861 G 861

RESULT 7
ID ATS9 HUMAN STANDARD; PRT; 1629 AA.
AC Q9P2N4; Q9NR29;
DT 16-OCT-2001 (Rel. 40, Created)
DE 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-9 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS 9) (ADAM-TS 9).
GN ADAMTS9 OR KIAA1312.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Fetal;
RX MEDLINE=20396138; PubMed=10936055;
RA Clark M.E., Kelnner G.S., Turbeville L.A., Boyer A., Arden K.A.,
RT "ADAMTS 9, a novel member of the ADAM-TS/Metallospodin gene
RT family.";
RL Genomics 67:343-350(2000).
RN [2]
RP SEQUENCE OF 159-1629 FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q9P2N4-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q9P2N4-2; Sequence=VSP_005499, VSP_005500;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.
CC EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,
CC PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN
CC COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR
CC THYMUS.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
CC SIMILARITY).
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FUIN ENDOPEPTIDASE (BY

```

```

CC SIMILARITY).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 12 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF261918; AAF89106.1; -
CC EMBL; AB037733; BAA92550.1; -
CC HSSP; P15167; IATL.
CC MEROPS; M12.021; -.
CC Genew; HGNC:13202; ADAMTS9.
CC MIM; 605421; -
CC GO; GO:0008181; P:tumor suppressor; TAS.
CC GO; GO:0007275; P:development; TAS.
CC GO; GO:0006516; P:glycoprotein catabolism; TAS.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR002870; Pep_M12B_propep.
CC InterPro; IPR001590; Repolysin.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR006025; Zn_Mpeptdse.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Repolysin; 1.
CC Pfam; PF00090; tsp_1; 11.
CC SMART; SM00209; TSP1; 12.
CC PROSITE; PS50215; ADAM_MERO; 1.
CC PROSITE; PS0546; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE; PS0427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS50092; TSP1; 12.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Extracellular matrix; Alternative splicing.
CC SIGNAL 1 18
CC PROPEP 19 287
CC CHAIN 288 1629
CC DOMAIN 293 499
CC DOMAIN 509 587
CC DOMAIN 588 643
CC DOMAIN 878 936
CC DOMAIN 645 752
CC DOMAIN 753 880
CC DOMAIN 997 1054
CC DOMAIN 1055 1104
CC DOMAIN 1107 1166
CC DOMAIN 1182 1240
CC DOMAIN 1241 1296
CC DOMAIN 1328 1379
CC DOMAIN 1382 1440
CC DOMAIN 1441 1494
CC DOMAIN 1497 1555
CC DOMAIN 1556 1613
CC DOMAIN 88 96
CC SITE 223 223
CC METAL 434 434
CC ACT SITE 435 435
CC METAL 438 438
CC METAL 444 444
CC CARBOHYD 112 112
CC CARBOHYD 135 135
CC CARBOHYD 271 271
CC CARBOHYD 749 749
CC CARBOHYD 840 840
CC CARBOHYD 1213 1213
CC CARBOHYD 1267 1267
CC VALSPIC 1064 1072
CC CLVTGKGH -> VRNEGCTFP (in isoform
CC Short).

```


FT VARSPLIC 1073 1629 /FTID=VSP_005499.
 FT Missing (in isoform Short).
 FT /FTID=VSP_005500.
 FT CONFLICT 367 367 P -> L (IN REF. 1).
 SQ SEQUENCE 1629 AA; 182649 MW; C1C4CEFF588941F CRC64;

Query Match 20.38; Score 1061.5; DB 1; Length 1629;
 Best Local Similarity 27.68; Pred. No. 3.3e-69;
 Matches 280; Conservative 148; Mismatches 334; Indels 253; Gaps 33;

Qy 100 RDMDEAGNHSQQTNTGTENQTLH-----VLTQDLVSAVEYDHRGDYVSHIMHQ 151
 Db 15 RDLAEMG---SPDAAAVRKRDLHPRQVKLETLGEVEIVSPIRVNALEGEPPPTNV-HFK 70
 Qy 152 RRRRAV-----AVSEVSLHRLGPRHDFHMDLRTSSLSVAPGFIQVTLG 197
 Db 71 RTRRSINSATDPWPAFASSSSSTSSQAHYLSAFQCGQFLNLTANAGFIAPLFTVTLG 130
 Qy 198 KTGKTSQVTLPPED-----FCFYQGLSLSHNSP-----SHGKFC----- 233
 Db 131 TPGVNTQKFKYSEAEELKHCFYKYGVNTNSENHTAVISLCSGLGTFRSHDGDYFISPLQS 190
 Qy 234 -----EGSTRTLKLCNSQKCP--DSVDFRAAQCAEHNRS-----RPRGHYKWKYTOV 281
 Db 191 MDEQDEEEQKPHIYRRSAPQREPSTGRHACDITSEHKNRSHDKKTKTARKWGERINL 250
 Qy 282 EADLCKLYCIAEGDFFFLSLNKVKDGTGPPCEDSRNVCIDGICLSVSTSAHMPQPKKE 341
 Db 251 AGDVAALN-----SGLATEAFSAVGNKT 273
 Qy 342 DLFLPDSYKSLRHSLSLRSHR--NEELNVEVLVVVKMMQNGHENITTYVLTILN 399
 Db 274 D-----NTRKTRTKRFLSPYRFEVVLVADNRVSYNG-ENLQHYILTMS 322
 Qy 400 MVSALFDGLMCK-----SGLATEAFSAVGNKT 273
 Db 323 IVASIKDPSIGNLINIVNLVIHNEQDPSISFNAQTLLKNCQWQSHKMSPGIHR 382
 Qy 418 DHAILLTGDLICSNKBPCTDLGPAPISGMCKYRSTCTINEDTGLGLAFTIAHESHNFG 477
 Db 383 DTAULLTQDICTR-AHDKCDTGLAELGTICDPYRSCSISDSGLSTAFTHAELGHVFN 441
 Qy 478 MHGEGNMCK-----KSEGNIMSTLAGRNGVPSWSPCSQYLHKLSTAAQCLADQPK 533
 Db 442 MPHD-DNNKCKEGBQSPQHVMAPTLNFYTPNPMWSKSRKYITEFLDTGYGCLLNEPE 500
 Qy 534 PVKEYKPEKLPGLBELYDANTCKWQFGEMAKLMLDPKDKDICALMCHRG--RKCETK 590
 Db 501 S-RPYPLPVQLPGLLYNNVKQCELI FPGSQVCPYMWQ---CRLMNCNVNGVHKGRITQ 556
 Qy 591 FMPAAGTICGHDWRCGQCVKYGDGPKETHGNSDWSWSPCSRTCCGGVSHSRSLC 650
 Db 557 HTPWADGTECFPGKHCKYGFVPMKMDVP-VTDGSGWSGSPFGFCSTCGGKTAIREC 615
 Qy 651 TNPKPSHGKFCGEGSTRTLKLCNSOKPRDSVDFFRAAQCAEHNRSRRFRGRHY----- 702
 Db 616 NRPEKNGKVCGERMKFSCNTEPCLKOKEDFDEQCA-----HFGKHFNGLNLPN 670
 Qy 703 -KWKP-YTQVEDQDLCKLYCIAEGDFFFLSLNKVKDGTGPPCEDSRNVCIDGICERVGCD 760
 Db 671 VRWPKYSGLMKDKRCLFCRVAGNTAYQLDRVIDGTGPGQDTNDCVQGLCRQAGCD 730
 Qy 761 NVLGSDAVEDVCGVCGNNSACTIHRGLYTKHHHTNOYHYHMTIPSGARSIYEMNVS- 819
 Db 731 HVLSKARDKCGVCGGNSCKTKVAGTFNTVH--GUNTIVRIPAGATIDVQRHSFSG 788
 Qy 820 -----TSYISVRNARYLNGHWTVDMFGR-YKFSGTTFDPRRSYNFENLIATGPTNET 874
 Db 789 ETTDDNNLALSSKGEFLNGFVVMKREIRIGNAVVEYSGSETAVERINSRDEQEQE 848
 Qy 875 LIVELLPOGR--NRCVANEYSMPRLGTEKQP----- 904
 Db 849 LLLQVLVSKLYNPDPVRSFNP-----IEDKFCQFVNSHGPNQACSKPCQGERKRLVCT 905

Qy 905 -----AOPSY-----TWAIIV-RSECSVSCG-GGRCLPV 930
 Db 906 RESQLTVSDQRDLRFPQGHTEPCGTCDLRHVASRSECSAQCGLGYRTLDI 960

RESULT 8

AT20 HUMAN STANDARD; PRT; 1911 AA.
 AC P59510;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE ADAMTS-20 precursor (EC 3.4.24 -) (A disintegrin and metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-TS20).
 GN ADAMTS20.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]_TaxID=9606;
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22513925; PubMed=12514189;
 RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M., Evanko S., Wight T.N., Leduc R., Apte S.S.;
 RA "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS subfamily related to Caenorhabditis elegans GON-1";
 RT J. Biol. Chem. 278:9503-9513(2003).
 RL [2]
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP TISSUE=Liver;
 RC Llanazares M., Cal S., Lopez-Otin C.;
 RT "Identification and characterization of ADAMTS-20 defines a novel subfamily of metalloproteinases-disintegrins with multiple thrombospondin-1 repeats and a unique GON-domain";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May play a role in tissue remodeling process occurring in both normal and pathological conditions.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P59510-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P59510-2; Sequence=VSP_007106, VSP_007107, VSP_007108;
 CC -!- TISSUE SPECIFICITY: Very sparingly expressed, although is detected at low levels in testis, prostate, ovary, heart, placenta, lung and pancreas. Overexpressed in several brain, colon and breast carcinomas.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 1 GON domain.
 CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; AF48804; AAO15766.1; -;
 CC EMBL; AJ515153; CAD56159.3; -;
 CC EMBL; AJ515154; CAD56160.2; -;
 CC SMART; SM00209; TSP1; 12.
 CC PROSITE; PS00215; ADAM_MEPHO; 1.
 CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

GN ADAMTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RN SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=98110583; PubMed=9441751;
RA Kuno K., Lizasa H., Ohno S., Matsushima K.,
RT "The exon/intron organization and chromosomal mapping of the mouse
RT ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.",
RL Genomics 46:466-471(1997).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=97150761; PubMed=8995297;
RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,
RA Matsushima K.,
RT "Molecular cloning of a gene encoding a new type of metalloproteinase-
RT disintegrin family protein with thrombospondin motifs as an
RT inflammation associated gene.",
RL J. Biol. Chem. 272:556-562(1997).
RN [3]
RN CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.
RX MEDLINE=99303657; PubMed=10373500;
RA Kuno K., Terashima Y., Matsushima K.,
RT "ADAMTS-1 is an active metalloproteinase associated with the
RT extracellular matrix.",
RL J. Biol. Chem. 274:18821-18826(1999).
RN [4]
RN FUNCTION.
RX MEDLINE=20389568; PubMed=10930576;
RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,
RA Ohno H., Matsushima K.,
RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.",
RL FEBS Lett. 478:241-245(2000).
RN [5]
RN FUNCTION, AND INDUCTION.
RX MEDLINE=20243757; PubMed=10781075;
RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
RA Richards J.S.,
RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and
RT cathepsin L proteases.",
RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
CC -!- FUNCTION: CLEAVES AGGECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-|-Leu-1692
CC site, within the chondroitin sulfate attachment domain.
CC -!- COPACITOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX.
CC -!- INDUCTION: INDUCED IN VITRO IN COLON ADENOCARCINOMA CELLS BY
CC INTERLEUKIN-1, OR IN VIVO IN KIDNEY AND HEART BY
CC LIPOLYSACCHARIDE. ALSO INDUCED BY LH STIMULATION IN GRANULOSA
CC CELLS OF PREOVULATORY FOLLICLES.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 7.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB001735; BAA24501.1; ALT INIT.
CC EMBL; D67076; BAA11088.1; ALT_FRAME.
DR MEROPS: M12.222;
DR MGD; MGI:109249; Adamts1.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR006025; Zn_M12peptidse.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; TSP_1; 3.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS0092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 48
FT PROPEP 49 253
FT CHAIN 254 968
FT DOMAIN 254 476
FT DOMAIN 477 559
FT DOMAIN 560 615
FT DOMAIN 618 725
FT DOMAIN 726 850
FT DOMAIN 855 911
FT DOMAIN 912 968
FT DOMAIN 135 199
FT SITE 206 206
FT METAL 402 402
FT ACT_SITE 403 403
FT METAL 406 406
FT METAL 412 412
FT CARBOHYD 548 548
FT CARBOHYD 721 721
FT CARBOHYD 765 765
FT CARBOHYD 783 783
FT CARBOHYD 946 946
FT MUTAGEN 403 403
FT CONFLICT 335 335
FT CONFLICT 425 425
SQ SEQUENCE 968 AA; 105841 MW; 42EBDAS5499FB6C1 CRC64;
Query Match 19.5%; Score 1022.5; DB 1; Length 968;
Best Local Similarity 30.6%; Pred. No. 1.2e-66;
Matches 266; Conservative 99; Mismatches 319; Indels 185; Gaps 32;
QY 167 LRLKPRHDFHMDLTSSSLVAPGHIVQTLQKTKSVQTLPPD---FCYQSSLSHR 223
DB 81 LRLDFAQQQLHLKLPQDPSGLAPGFTLQTVGRSPSEAHLDPTGDLAHCFYSGTVANGD- 139
QY 224 NSPSHGKFCGSTRLLKLCNSQKPRDSVDFRAAQCAEHNSRRFRGRHYKWKPYQVEA 283
DB 140 -----FGSAARLSLCEG-----VRGAFY----- 157
QY 284 DLCKLYCIABGDFPFFS-----LSNKVKDGTGPPCEDSRN-----VVDIG 322
DB 158 -----LQGEFFIQAPGAVATERLAPAVEBESSARFQPHILRRRRRGGGAKC--G 207
QY 323 ICESLVSVSTSAHMQ-----PPKEDLFLPDVYKSLRHSKLSLRSHRNEELN 370
DB 208 VMDDELTLTSRPSRPSQNRQWVRDPTQDAG--KPSGPGSIRKGRFV-----SSPRY 260
QY 371 VETLVVDKGMQNHGHENITTYVLTILNMVSALFK----- 406

Db 261 VETMLVADQSMADFHG-SGLKHVYLLTFLSVAAREFYKHSIRNSISLVVVKILVIVEEQK 319
 Qy 407 -----DGLMKDGTGRHDHAILLGLDSCWKNPCCPTLGFAPIS 445
 Db 320 PEVTSNAALFLRNCWQKOHNSPDRDEHYDTAILFTRQDLG--SHTCDTLGMADV 377
 Qy 446 GMSKYSRSCVINEDTGLGFLAFTIAHSGHNFHGDGEGNMCKKSEG-----NIMSPTIA 500
 Db 378 TVCDPSRSCSVIEDGLQAFHTTAHELGHVFNPHD-DAKHCASLANGVTGDSHLMASMLS 436
 Qy 501 GRNGVFSWSPCSQYHLKFLTAQAICLADQPPVKEYKPEKLPGLIYDANTOCKWQFG 560
 Db 437 SLDSQWSPSCSYMYVTSFLDNGHGBCLMD--KPQNPILPSLPGLTLYDANRQCOFTFG 494
 Qy 561 EKAKLMLDPKDKICAKLCHGRIGRK-----CETKMPAAAGFTICGDMWCRGGGCVKYG 615
 Db 495 EESKHC--PDAATCTTLAC--TGTSGLLVCTQKHPFWDGTSCGEGKCVSGKCVNKT 550
 Qy 616 D--EGPPTTHGWSWSPCSRTCGGVSHRSRLCTNPKPSHGKFCGEGSTRITKLON 673
 Db 551 DMKHFATPVHSGMPGFWGDCSRTCGGVQYTMRECDNPPVKNKGKFCGKRVRYRSCN 610
 Qy 674 SOKCP-RDSVDFRAAQAENH--SRPRGRH--YKWKPE-YTQVEDDCLKLYCIAEGPFD 727
 Db 611 IEDCPDNNGKTFREOCEAHNEFSKASFGNEPTVETPKYAGSPKDRCKLICEAKIGY 670
 Qy 728 FFSLSNKKVKGTPCSDSRNVICDIGICERVCGDNVLGSDAIVEDVCGVGNNSACTIHRG 787
 Db 671 FVLQPKVVDGTPCSPDSTVCGVQCVKAGCDRIIDSKKKFKDCKGCGGNGSTCKMWSG 730
 Qy 788 LYTKHHTNQYH-MVTIPSGARSIRIYEMNV-----STSYISVRNALRYLYLNGHWTVD 841
 Db 731 IVTS---TRPGYHIVIPAGATNIEVHRNQRSGRNGSFLAIRAADGTLYLNGFTLS 787
 Qy 842 WPCR-YKFSGTTFDYRSYNEPENLIATGPTNETLIVELFQGR--NPGVANEYSMPRLG 898
 Db 788 TLEQULYKGVLYSGSSAALAIRSFSLKEPLATIQVLNWHGHALRPKIKFTYEMKK-- 845
 Qy 899 TEKQPPQPSYT-WAIVR-SECSVSCGG 925
 Db 846 KTESFNAIPFSEWIEWGECSKTCGSG 874

RESULT 10

ATSI RAT STANDARD; PRO; 967 AA.
 AC QPNUQL: QSER1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
 GN ADAMTS1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=Brain;
 RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
 RA Little S.P.;
 RT "Induction of a disintegrin and metalloproteinase with the
 RT thrombospondin type 1 motif (ADAMTS).";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 18-967 FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=20304099; PubMed=10847486;
 RA Diamantis I., Luetli M., Hoelsli M., Reichen J.;
 RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
 RT endothelial cells in cirrhotic rats.";
 RL Liver 20:165-172(2000).

CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
 CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
 CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
 CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
 CC SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1693-Glu-I-Leu-1684
 CC site, within the chondroitin sulfate attachment domain.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -!- INDUCTION: DOWN-REGULATED IN ENDOTHELIAL CELLS DERIVED FROM
 CC CIRRHOTIC LIVER.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF149118; AAC34012.1; .
 CC EMBL: AF304446; AAG29823.1; .
 CC MEROPS: M12_222; .
 CC InterPro: IPR004586; ADAM_cysteine.
 CC InterPro: IPR001762; Disintegrin.
 CC InterPro: IPR002870; Pep_M12B_propep.
 CC InterPro: IPR001590; Reprolysin.
 CC InterPro: IPR000884; TSP1.
 CC InterPro: IPR006025; Zn_Mtpeptidse.
 CC Pfam: PF01562; Pep_M12B_propep; 1.
 CC Pfam: PF01421; Reprolysin; 1.
 CC Pfam: PF00090; csp; 3.
 CC SMART: SM00608; ACR; 1.
 CC SMART: SM00209; TSP1; 3.
 CC PROSITE: PS0215; ADAM_MEPPO; 1.
 CC PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 CC PROSITE: PS0214; DISINTEGRIN_2; FALSE_NEG.
 CC PROSITE: PS00142; ZINC_PROTEASE; 1.
 CC PROSITE: PS00992; TSP1; 3.
 CC Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
 CC Repeat; Extracellular matrix; Heparin-binding.
 CC SIGNAL 1 54
 CC PROPEP 55 252
 CC CHAIN 253 967
 CC ADAMTS-1.
 CC METALLOPROTEASE.
 CC DISINTEGRIN-LIKE.
 CC TSP TYPE-1 1.
 CC CYS-RICH.
 CC SPACER.
 CC TSP TYPE-1 2.
 CC TSP TYPE-1 3.
 CC POLY-ARG.
 CC CYSTEINE SWITCH (POTENTIAL).
 CC ZINC (CATALYTIC) (BY SIMILARITY).
 CC ZINC (CATALYTIC) (BY SIMILARITY).
 CC ZINC (CATALYTIC) (BY SIMILARITY).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC I -> V (IN REF. 2).
 CC KFRSSQ -> RSRGSL (IN REF. 2).
 CC CONFLICT 26
 CC CONFLICT 31


```

FT CONFLICT 49 49 V -> A (IN REF. 2).
FT CONFLICT 72 72 R -> P (IN REF. 2).
FT CONFLICT 79 79 L -> TR (IN REF. 2).
FT CONFLICT 249 249 R -> G (IN REF. 2).
FT CONFLICT 262 265 TMLV -> NLTK (IN REF. 2).
FT CONFLICT 607 607 S -> F (IN REF. 2).
FT CONFLICT 936 936 L -> V (IN REF. 2).
FT CONFLICT 962 962 Y -> T (IN REF. 2).
SQ SEQUENCE 967 AA; 105705 MW; F93C864F6DCB4CF CRC64;

Query Match
Best Local Similarity 19.5%; Score 1021.5; DB 1; Length 967;
Matches 269; Conservativity 93; Mismatches 304; Indels 221; Gaps 32;

QY 167 LRLKPRHDFHMDLRTSSSLVAPGFIQTGLTKTKSVQTLPPED---FCFYQGSLSHR 223
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 80 LRLDAPGQQLKLPDGSGLAPGFTLTQVGRSPGSAQHLDPGTGLARCFYSGTV---N 136
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 224 NSPSHGGRFCGSTRTLKLCNSQKCPDSVDFAAQAHAHNRFRGRHYKWKPYTQVEA 283
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 137 GDPS-----SAALSCEG-----VRGAFY-----156

QY 284 DLCKLYTAGDFDFFLSLNKVDGTPCSDSRNVCIDGICELSVSTSAHMPQPKEDL 343
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 157 -----LQGEFF-----IQAPAVATERLVAEPKEE-183

QY 344 FILDEYKSLRHKR-----SLLRSHRNEELN-----370

DB 184 SIAPPRFHLRRRGSGAGKCGWDETLFTSNGRESQNTPCWPLNPTPGACKPT 243

QY 371 -----VETLVVKKMMQNHENITTYVLTILNMVSALEK-----406
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 244 GPGSIRKRPVSSPRVETMLVADQSMADPHG-SGLKHYLTLSFVAARFYKHPISRNSI 302
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 407 -----DELGKGTGRDHAILLTGIDIC 429
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 303 SLUVVKVILVIEEQGPVTSNAALTANFCSQKQNSPSDRDPEHYDTAILFTEDQLC 362

QY 430 SWNEPFCPTLGPAPISGCMKYRSTINEDTGLGLAFTIAHSGHNFMIHDGNNCKK 489
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 363 G-SHTCDTLGMADVGTGDCPSRSCSVIEDDGLQAAFTTAHELGHVENMPHD-DAKICAS 419

QY 490 SEG-----NINSPTLAGNVFSWSPCSQRYLHLFLSTAQAICLAQPKPKVEKYPEKL 544
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 420 FNGVSGDHLWASMLUSSJHSPQSPSCAYMVTSTFLONGHGECLMD--KPNQPKLPSDL 477

QY 545 PCELYDANTQCKWQGEKAKLMDLFFKDKICKALWCHRIGRK-----CETKFMFAARGTI 599
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 478 FGLTYDANRQCFTEGEESTHC--PDAASTCTSLWC--TGRSGLLVCQTKHFPWADGTS 533

QY 600 CGHDNWRGGQCKVYGD--EGPKPTHGWSHNSPSCSRCCGGVSHRSLCTNPKPSH 657
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 534 CGEGKWCYSGKCVNKTDMKGFATPVHGSWGPWPGDCSRTCCGGVQYTWRECDNPVFN 593

QY 658 GSKFCBGSSTRTLKLCNSQKCP-RDSVDFRAAQCAEHN--SRRFRGRH--YKWKP-YTQVE 711
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 594 GKCYCEKVRVRSNIEDCPDNNGKTFREQCEAHNEFSKASFGNEPTVENTKYAGVS 653

QY 712 DDCLKLYCIAGDFPFLSLNVDGTPCSDSRNVCIDGICERVCNDVULGSDAEDV 771
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 654 PDORCKLTCEAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQCGKAGCDRIIDKKRFDK 713

QY 772 CGVCGNNSACTIHRGLYTKHHTNQYH-VYTIPIGARSIRIVEMNV-----STSYISV 825
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 714 CGVCGNGSTCKKISGTVTS-----TRPGHDIVTIPAGATNIEVHRNPRSGNNGSLAI 770

QY 826 RNALARYLNGHWTWDFQR-YKFGSTTFDYRRSYNEPENIATGPTNETLIVELLFQGR 884
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 771 RAADGTIYILNGNFTLSTLQDLTYKGTVLVRSKSSAALERIRSPLKEPLTIQVLWVG 830

QY 885 --NPGVANEYSNPLGTGEKQP--NQPSYT-WAIVR-SECVSCGG 925
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 831 ALRPKIKYTFMRK-----KTEPFNAIPTFSEVIEWGECKSKTCGG 873

```

RESULT 11

AT19 HUMAN

```

ID AT19 HUMAN STANDARD; PRT; 1207 AA.
AC Q8TE59;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ADAMTS-19 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 19) (ADAM-TS 19) (ADAM-TS19).
GN ADAMTS19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-i domains.";
RL Gene 283:49-62(2002).
CC -! COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -! SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -! TISSUE SPECIFICITY: Expressed in fetal lung, but not in any adult
CC tissues examined. Expression was detected in an osteosarcoma CDNA
CC library.
CC -! PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -! SIMILARITY: Belongs to peptidase family M128.
CC -! SIMILARITY: Contains 1 disintegrin-like domain.
CC -! SIMILARITY: Contains 1 PLAC domain.
CC -! SIMILARITY: Contains 5 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AJ311904; CAC84565.1; -.
CC Genew; HGNC:17111; ADAMTS19.
CC MIM; 607513; -.
CC InterPro; IPR002870; Pep_M12B_propep.
CC InterPro; IPR001590; Reprolysin.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR006025; Zn_Mtpeptdse.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC SMART; SM00090; TSP_1; 5.
CC SMART; SM00209; TSP1; 5.
CC PROSITE; PS50215; ADAM_NEPRO; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS50092; TSP1; 5.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 316 BY SIMILARITY.
FT CHAIN 317 1207 ADAMTS-19.
FT DOMAIN 317 545 METALLOPROTEASE.
FT DOMAIN 546 633 DISINTEGRIN-LIKE.
FT DOMAIN 634 686 TSP TYPE-1 1.
FT DOMAIN 687 790 CIS-RICH.
FT DOMAIN 791 914 SPACER.

```


CC CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC CC -!- SIMILARITY: Belongs to peptidase family M12B.

CC CC -!- SIMILARITY: Contains 1 disintegrin-like domain.

CC CC -!- SIMILARITY: Contains 5 TSP type-1 domains.

CC CC -!- CAUTION: By homology with the human sequence, it is uncertain whether Met-1 or Met-5 is the initiator.

CC CC -----

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC -----

CC CC EMBL; AY135183; AAM10155.1; ..

CC CC MGD; MG1:2442875; Adamts19.

CC CC SMART; SM00608; ACR; 1.

CC CC SMART; SM00209; TSP1; 5.

CC CC PROSITE; PS0215; ADAM MEPRO; 1.

CC CC PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.

CC CC PROSITE; PS00427; DISINTEGRIN 1; FALSE NEG.

CC CC PROSITE; PS0214; DISINTEGRIN 2; FALSE NEG.

CC CC PROSITE; PS0092; TSP1; 5.

CC CC PROSITE; PS0142; ZINC PROTEASE; 1.

CC CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Repeat; Extracellular matrix.

CC CC SIGNAL 1 30 POTENTIAL.

CC CC PROPEP 31 319 BY SIMILARITY.

CC CC CHAIN 320 1210 ADAMTS-19.

CC CC DOMAIN 320 548 METALLOPROTEASE.

CC CC DOMAIN 548 636 DISINTEGRIN-LIKE.

CC CC DOMAIN 637 689 TSP TYPE-1 1.

CC CC DOMAIN 690 793 CYS-RICH.

CC CC DOMAIN 794 917 SPACER.

CC CC DOMAIN 918 978 TSP TYPE-1 2.

CC CC DOMAIN 979 1040 TSP TYPE-1 3.

CC CC DOMAIN 1042 1086 TSP TYPE-1 4.

CC CC DOMAIN 1090 1147 TSP TYPE-1 5.

CC CC DOMAIN 1167 1198 PLAC.

CC CC SITE 297 297 CYSTEINE SWITCH (POTENTIAL).

CC CC METAL 485 485 ZINC (CATALYTIC) (BY SIMILARITY).

CC CC ACT_SITE 486 486 ZINC (CATALYTIC) (BY SIMILARITY).

CC CC METAL 489 489 ZINC (CATALYTIC) (BY SIMILARITY).

CC CC METAL 495 495 ZINC (CATALYTIC) (BY SIMILARITY).

CC CC CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC CARBOHYD 800 800 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC CARBOHYD 910 910 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC CARBOHYD 931 931 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC CARBOHYD 952 952 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC CARBOHYD 1012 1012 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC SEQUENCE 1210 AA; 134560 MW; 0A8812ABAB4BB7A2 CRC64;

Query Match 18.9%; Score 989; DB 1; Length 1210;

Best Local Similarity 28.4%; Pred. No. 4.3e-64;

Matches 295; Conservative 133; Mismatches 362; Indels 250; Gaps 39;

QY 1 MKPRARGRGL-AALMLLAQVAEQVS-----PGRSHQGRNGSGGLEAS----- 44

DB 33 LTPDLSEWVFPALWRRESLNATGLSGSSDPGSRSSGGGRGQASGSREVSVARA 92

QY 45 PRLLSRGPRRLTAMGFLPSAGTCVVRHTRSGSANEPEPASSTRGAAGLDGKGRDME 104

DB 93 PQEATRGQSEWFGSP-----EPGAEDDEELESQELPRGSSG-----DT 133

QY 105 A---GNHRSQQTNGTENQTLHLVLTQYDLVSAYEVDRGDIYSHEIMHQRRAVAVSE 161

DB 134 ALSSGTPASQPLPQRPSPSPPPAQOQEEPSAEV----- 168

QY 162 VESLHLRLKGRPHDFHMDLRTSSSLVAPGFIVQTILGKTG---TKSV---QTLPPEDRCF 214

DB 169 -----LLRIPALSRDLYLLLRDRGRFLAQRFAVEQWPKPGDPDPTRATADPGSSLLPDASCF 224

QY 215 YQGSLSHRNSPSH-----GG-----KFCGSTRITLKLNSQKCPRDSVDFAA 258

DB 225 YTGTVLRHPGSLASFSTCGGLMGFTQLNEDFLFIEPNDTMAILG----- 270

QY 259 QCAEHNSRFRGRHYKWKPYTQVEADLCKLYCIAEGDFFFSLSNVKVKGDTPCSESRNV 318

DB 271 -----HPRLYRQR-----STEEKV---TENSAYRRH 296

QY 319 CIDGICELSVSTSAHMPQPKEDFILPDEYKSLRHRSLRSHR-NEELNVTLVVV 377

DB 297 C-----GVISDKGR---PRS-----KKTADNRREKRYSKLSQEIYETVVVA 336

QY 378 DKKMQNHGHENITTVTLTKNVKSALEFKDGLM----- 410

DB 337 DPAMVSYHGADAARRFILTILNVFNLFQKSLGVQVNLRLVKLILLHETPADLYIGHHG 396

QY 411 -----GKQTRHDHAILLTGLDICSWKNEPCDTLGF 441

DB 397 EKULESFCWKQHEEFGRNDVHLEMTSWGEDIAAADAAILITRKDFCVHKDEPCDTVG 456

QY 442 APISGMSKYRSCITNEDTGLAFTIAHSGHNFMIHDGSGNCKKSEG-NIKS-PTL 499

DB 457 AVLNGWCSEKRCITAEENGLNLAFTIAHEMGNHGINHNDHPSC--ADGLHMSGEW 514

QY 500 AGRN-GVFSWSPCSQYLLHKFLSTAQAICLA-DQPKVPKBYKYPEKLPGELYDANTQCKM 557

DB 515 KQNLGDVSWSRCSKEDLERFLRSKASSCLLHTDPQSLSSVLVPSKLPGMAYTADQCQ 574

QY 558 QFGEKAKLMDPKDKDICALWCHRIGR-KCETKMPAAEAGTICGHDMMCRGGQCVKYGD 616

DB 575 LFGPIASFQ-EMQHVICTGLWCKVEGEAECKTLQPPMDGTDGDFGRKCKAGECTR--- 630

QY 617 EGPKEPTH--GHWSWSSWSPCSRTCGGVSHSRSLCTNPKPSHGKFCGSTRITLKLNS 674

DB 631 RTPAPEHLAGEWSPSS---CSFSCSSGVSSSRKRC--PGLGSEARDCNGPRKQYRICEN 685

QY 675 QKCPRDSVDFAAQCAEHNSRFRGRH-YKWKPYTQVEDQDLCKLYCIAEGDFFFSLSN 733

DB 686 PPCPAGLPGFRDWQCOAYSVRTSYPKHALOMQ--AVFDEEKPCLALFCSVPVGEQVLLSE 743

QY 734 KYKQGTGPGSEDSRNVICDIGICERVCNVLGSDADEVDCVCGNGNSACTIHRGLYTKHH 793

DB 744 KWDGTSQYGLDICANGRCQKACDGLLSLAREDHCGVCGNGKSCVKIKGDF-NHT 802

QY 794 HTNQYHMTIPSGARSIRIYEMNVSTSYISVRNALRYRYLNGHWTVDMWPGRYKFSGITP 853

DB 803 RGAGYVEVLVIPAGARRIKVVEEKPAHSFLALRDASKQ-SINSDWKIEHSGAFSLAGITV 861

QY 854 DYRRSYNEPENLIATGPTNETL-IVELLFOGPNPGVAWEYSMPRLGTGKQPPAQ----- 906

DB 862 HYLRR-GLWEKISAGPPTTTLHLVLVLLFQDQNYGLHYEYTVP---SDPLPDNQSSKEFG 917

QY 907 PSYTWALVR-SECSVSCGGG 925

DB 918 PLEMTWTHAGWGDNATCGG 937

RESULT 14

AT17 HUMAN

ID AT17 HUMAN STANDARD; PRT; 1095 AA.

AC Q8TE56;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE ADAMTS-17 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 17) (ADAM-TS 17) (ADAM-TS17).

GN ADAMTS17.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1] SEQUENCE FROM N.A.
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
 Lopez-Otin C.;
 RT "Cloning, expression analysis, and structural characterization of
 seven novel human ADAMTSs, a family of metalloproteinases with
 disintegrin and thrombospondin-1 domains.";
 RL Gene 283:49-62(2002).
 CC -!- SUBCELLULAR LOCATION: secreted. Associated with the extracellular
 matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in fetal lung, in adult brain,
 prostate, and liver.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 1 PLAC domain.
 CC -!- SIMILARITY: Contains 5 TSP type-1 domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; AJ315735; CAC86016.1; -.
 DR MEROPS; M12.027; -.
 DR Genew; HGNC:17109; ADAMTS17.
 DR MIM; 607511;
 DR InterPro; IPR002870; Pep_M12B_propep.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR006025; Zn_MTPeptide.
 DR InterPro; IPR007087; Zn_C2H2.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; TSP1; 5.
 DR SMART; SMO0209; TSP1; 5.
 DR PROSITE; PS02015; ADAM_MEPRO; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS02114; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE; PS00092; TSP1; 5.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix.
 FT SIGNAL 1 27 POTENTIAL.
 FT PROPEP 28 223 BY SIMILARITY.
 FT CHAIN 224 1095 ADAMTS-17.
 FT DOMAIN 224 452 METALLOPROTEASE.
 FT DOMAIN 453 542 DISINTEGRIN-LIKE.
 FT DOMAIN 543 598 TSP TYPE-1 1.
 FT DOMAIN 599 701 CYS-RICH.
 FT DOMAIN 702 779 SPACER.
 FT DOMAIN 800 860 TSP TYPE-1 2.
 FT DOMAIN 861 922 TSP TYPE-1 3.
 FT DOMAIN 925 968 TSP TYPE-1 4.
 FT DOMAIN 972 1029 TSP TYPE-1 5.
 FT DOMAIN 1049 1080 PLAC.
 FT DOMAIN 60 120 ARG-RICH.
 FT SITE 201 201 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 389 389 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 390 390 BY SIMILARITY.
 FT METAL 393 393 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 399 399 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 785 785 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 790 790 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 832 832 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 839 839 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 894 894 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1095 AA; 121099 MW; A5C6540484754D5D CRC64;
 Query Match 18.8%; Score 985; DB 1; Length 1095;
 Best Local Similarity 30.3%; Pred. No. 7.5e-64;
 Matches 259; Conservative 103; Mismatches 281; Indels 212; Gaps 32;
 QY 165 LHLRLKGRPHDMLRTSSLSVAPGFIVOTLGTGKTSVQTLPPEDFCFQGSLSHRN 224
 DB 83 LHLPAFG--RDLYLQLRDLRFLSRGFVEEAGARRRG---RPAELCYSGVLGH-- 134
 QY 225 SPSHGKFCGEGTTLKLSNSQ-----KCPRDSVDFRAAQCAEHNRSRPRGRHY 273
 DB 135 -----PGSLVSLSACGAAGGLVLIQLGQEV---LIQPLNNSQGFSGREHLIR 182
 QY 274 KWK--PYTQVEA---DLCKLYCIAEGDFPFSSLNKVKD--GTPCEDSRNVICDIGE 325
 DB 183 KWSLTPSPSAQAORPEQLCKV-----LTEKKKPTWGP-----SRD----- 218
 QY 326 LSVVSTSAHMPPOPKEDLFLPDEYKSLRHRKSLRSLRNEELNVETLVVVKMMQNH 385
 DB 219 -----WRERRNAIR--LTSEHTVETLVVADAMVQYH 248
 QY 386 GHENITTYVILNNVSNALFKDGLMG----- 411
 DB 249 GAAEAQRFILTMNVMNMFQHSIGIKINIQVTKLVLLRQRPAAKLSIGHGERSLESFC 308
 QY 412 -----KDGTRH-----DHAILLTGLDICSWKNEPCDTLGFAPISQWCS 449
 DB 309 HWQNEEYGYARYLGNQVPGCKDDPLVDAAVFTRTDFCVHKDEPCTDVGIALGVCS 368
 QY 450 KYRSTINEDTGLAFTIAHESGHNFQIHDEGNMCKKSEGNIMS-PTLAGRN-GVFS 507
 DB 369 AKRKCVLAEDNGLNAFTIAHELGNLGMNDDHSSC-AGRSIMSGEWMKGNPNDLS 427
 QY 508 WSPCSRQYLHFLSTAQAIC-LADQPKPVKEYKYPEKLPGLYDANTQCKWQFCEKAKLC 566
 DB 428 WSSCSRDLLENFLSKVSTCLLVTPRSQHTVRLPHKLPGMHYANSCQQLFGMNATFC 487
 QY 567 MLDFKKDJCKALWCHRIG-RKCTKFMPPAEGTTCGHDMCRGCGQCVKYGDGPKPTH-- 623
 DB 488 R-NMHLMCAGLWCLVEGDTCKTKLDPFLDGTGCGADKMCRAGECV---SKTPIPEHVD 543
 QY 624 GHWSDSWSNCPSCRTCGGVSHRSRLCTNPKPSHGKFCGSTRTLKLSQKCPDSVD 683
 DB 544 GWSWPGAWSCSRICGARGFRKCONPPPGFGTHCPGASVEHAUCENLPCPKGLPS 603
 QY 684 FRAAQCAEHNRRFRGRHYKWKPYTQ-----VEDQDLCKLYCIAEGDFPFSSLNKVKD 737
 DB 604 FRDQCCQAH-----RLSPKKKGLLTAVVVDKPCELYCSPLGKESPLLVAADVLD 654
 QY 738 GTPCEDSRNVICIGECRVCNDVLGSDADEVGVCNCGNNSACTIHRGLYTKHHHTNQ 797
 DB 655 GTPCGPYETDLCVHGKCKQKICGDLIGSAKEDRCGVCSGDKTKCHLVKGF----- 706
 QY 798 YYHMTIPSGARSIRIYEMNVTSTYISVRNALRERYLNGHWTVDWPGRYKFSGTTFYRR 857
 DB 707 -----SHARGTALKDSGKS-----INSDWKIELPGEFQIAGTTVYVR 745
 QY 858 SYNEPENLIATGPTNETL-IVELLFQGRNPOVAWEYSMP--RLGTEKQPPAQS---YTW 911
 DB 746 R-GLWEKISAKGPTKLPPLHMLVLLFHQDYGIHYEYVFPVNRTAENQSEPEKQDSLF 804
 QY 912 AIVRSE-CVSVCGGG 925
 DB 805 THSGWEGCSVQCGGG 819
 RESULT 15
 ID ATSL HUMAN STANDARD; PRT; 967 AA.
 AC Q9UHI8; Q9NSJ8; Q9P2K0; Q9UH83; Q9UP80;

Search completed: October 28, 2003, 23:41:55
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 23:23:19 ; Search time 76 Seconds

(without alignments)
3232.448 Million cell updates/sec

Title: US-09-981-151A-8

Perfect score: 5236
Sequence: 1 MKPRAGRWGLAALWMLAQ.....LEAQCPSATAYIALAFLES 952

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_todant:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3858.5	73.7	1072	4 Q8TE57	Q8TE57 homo sapien
2	3400	64.9	900	11 Q8K206	Q8K206 mus musculus
3	3280	62.6	1021	4 Q8VE2	Q8VE2 homo sapien
4	2408	46.0	1081	4 Q8TE60	Q8TE60 homo sapien
5	2257	43.1	1092	11 Q8BKX1	Q8BKX1 mus musculus
6	1218	23.3	1009	11 Q8BKV1	Q8BKV1 mus musculus
7	1182	22.6	1070	11 Q8CG28	Q8CG28 mus musculus
8	1075	20.5	1059	5 Q8W493	Q8W493 drosophila
9	1018.5	19.5	1207	4 Q8TE59	Q8TE59 homo sapien
10	984.5	18.8	967	4 Q8NE26	Q8NE26 homo sapien
11	965.5	18.4	2165	5 Q19791	Q19791 caenorhabdi
12	944.5	18.0	759	6 Q8HZW8	Q8HZW8 equus cabal
13	873.5	16.7	497	11 Q8BZD1	Q8BZD1 mus musculus
14	872.5	16.7	833	11 Q8K384	Q8K384 mus musculus
15	872.5	16.7	845	11 Q8BNJ2	Q8BNJ2 mus musculus
16	861.5	16.5	769	5 Q8MRU5	Q8MRU5 drosophila

17	821.5	15.7	623	11	Q8BGP4	Q8BGP4 mus musculus
18	819.5	15.7	1688	5	Q8SX80	Q8SX80 drosophila
19	783	15.0	1427	4	Q96L37	Q96L37 homo sapien
20	745	14.2	820	11	Q8C9W3	Q8C9W3 mus musculus
21	640.5	12.2	1235	4	Q95428	Q95428 homo sapien
22	632.5	12.1	1280	11	Q9EPX2	Q9EPX2 mus musculus
23	617	11.8	2174	5	Q9GQR0	Q9GQR0 drosophila
24	615.5	11.8	3198	5	Q9U8G8	Q9U8G8 manduca sex
25	615	11.7	3060	5	Q9AV4	Q9AV4 drosophila
26	549	10.5	356	4	Q8IW95	Q8IW95 homo sapien
27	533	10.2	1091	5	Q9M126	Q9M126 drosophila
28	525.5	10.0	1444	5	Q17591	Q17591 caenorhabdi
29	525.5	10.0	1461	5	Q8MYA8	Q8MYA8 caenorhabdi
30	508.5	9.7	1487	5	Q8MPV5	Q8MPV5 caenorhabdi
31	508.5	9.7	1558	5	Q8I710	Q8I710 caenorhabdi
32	508.5	9.7	2167	5	Q76840	Q76840 caenorhabdi
33	502	9.6	269	6	Q9GL54	Q9GL54 oryctolagus
34	500.5	9.6	790	5	Q8MRE1	Q8MRE1 drosophila
35	499.5	9.5	951	4	Q60345	Q60345 homo sapien
36	499	9.5	790	5	Q8T458	Q8T458 drosophila
37	489.5	9.3	1572	5	Q44938	Q44938 haemochus
38	473.5	9.0	766	4	Q82987	Q82987 homo sapien
39	472.5	9.0	872	5	Q22580	Q22580 caenorhabdi
40	470.5	9.0	1055	5	Q19204	Q19204 caenorhabdi
41	461.5	8.8	1020	5	Q8IU50	Q8IU50 caenorhabdi
42	455.5	8.7	439	4	Q8N6G6	Q8N6G6 homo sapien
43	451.5	8.6	525	4	Q96R94	Q96R94 homo sapien
44	451.5	8.6	880	5	Q8MSF8	Q8MSF8 drosophila
45	449.5	8.6	377	11	Q8BLI0	Q8BLI0 mus musculus

ALIGNMENTS

RESULT 1

Q8TE57 PRELIMINARY; PRT; 1072 AA.
 ID Q8TE57
 AC Q8TE57;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Metalloprotease disintegrin 16 with thrombospondin type I motif.
 GN ADAMTS16.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
 RA Lopez-Otin C.;
 RT "Cloning, expression analysis, and structural characterization of
 RT seven novel human ADAMTSs, a family of metalloproteinases with
 RT disintegrin and thrombospondin-1 domains.";
 RL Gene 283:49-62 (2002).
 DR EMBL; AJ315734; CAC86015.1; --
 DR MEROPS; M12.026; --
 DR Genew; HGNC:17108; ADAMTS16.
 DR InterPro; IPR006130; Asp/Ori_Cotranf.
 DR InterPro; IPR002870; Pep_M12B_propep.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR SMART; SM00090; tsp_1; 3.
 DR SMART; SM00209; TSP1; 4.
 DR PROSITE; PS00215; ADAM_MEPRO; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
 DR PROSITE; PS00092; TSP1; 3.
 KW Integrin; Protease.
 SQ SEQUENCE 1072 AA; 119491 MW; 081EEFF78F47D061 CRC64;

Query Match 73.7%; Score 3858.5; DB 4; Length 1072;
 Best Local Similarity 75.6%; Pred. No. 0;
 Matches 740; Conservative 24; Mismatches 72; Indels 143; Gaps 13;

QY 1 MKPRARGWGLAALWMLLAQVAVQVSGRSHQNRGSGOLEASPPRLLSRGRRLTAMS 60
 DQ 1 MKPRARGWGLAALWMLLAQVAVQVSGRSHQNRGSGOLEASPPRLLSRGRRLTAMS 60
 QY 61 PLASAGTCVRHGTGSSAMPERPASSSTRAGAGLDGKGRMDDEAGNHRSGQTNTGTENQ 120
 DQ 32 PAAAA-----PGSPVFRPPPAERPG-----WNEKG-----58
 QY 121 TLAVLTQYDLVAYEDHGRDYVSHEIMHQRRAVAVSEVESLRLKLGPRHDFHMDL 180
 DQ 59 -----EYDLVAYEDHGRDYVSHEIMHQRRAVAVSEVESLRLKLGPRHDFHMDL 112
 QY 181 RTSSSLVAPGFIYVTLGKTGTSVOTLPEDFCYQSLRSHRNSPSHGKCEGSTRIL 240
 DQ 113 RTSSSLVAPGFIYVTLGKTGTSVOTLPEDFCYQSLRSHRNS-SVALSTCQGLSGMI 171
 QY 241 KLCNSQKCPDSDVDFRAAQCAEHSNRFRGR-----HYKWKPYTOVEA-DLCKLYCI 291
 DQ 172 R-----TEADYFLRPLPSHLWKL-GRAAQSGSPSHVLYKRSTEPHAPGASEVLVT 222
 QY 292 AEGFDFFPS--LSNKVKGDTFCSEDSNRVCIIDIGICELSVSTSAHNPQPKEDLFLPDE 349
 DQ 223 SRTWELAHQPLHSDRLGLF---QKHFC-----GRRKKYMPQPKEDLFLPDE 270
 QY 350 YKSLRHRKSLRSHRNEELVETLVVVDKMKMNHGHNITTYVLTILNMYSAFPK-- 407
 DQ 271 YKSLRHRKSLRSHRNEELVETLVVVDKMKMNHGHNITTYVLTILNMYSAFPKDT 330
 QY 408 -----GLMKDGTGRHDAILLTGID 427
 DQ 331 IGGNINIAIVGLILLEDEQGLVISHHADTLSTFCQWQSLGKMGKGTGRHDAILLTGID 390
 QY 428 ICSWKNEPCDTLGFAPISGCKYRSCTINEDTGLGFTIAHESGHNFNMHDSGNMC 487
 DQ 391 ICSWKNEPCDTLGFAPISGCKYRSCTINEDTGLGFTIAHESGHNFNMHDSGNMC 450
 QY 488 KKSSEGNIMSPFLAGNGVFSWSPCSROYLHKFLSTQAQICLADQPKVKEYKYPEKLPGE 547
 DQ 451 KKSSEGNIMSPFLAGNGVFSWSPCSROYLHKFLSTQAQICLADQPKVKEYKYPEKLPGE 510
 QY 548 LYDANTOCKWFGSKAKLMDLDFKKDICKALWCHRIKCKETKFMFAAEGTICGHDMWCR 607
 DQ 511 LYDANTOCKWFGSKAKLMDLDFKKDICKALWCHRIKCKETKFMFAAEGTICGHDMWCR 570
 QY 608 GGQCVKYGDEGPKPTHGHSWSSWSPCSRTCCGGVSHRSLCTNPKPSHGKPCGSTR 667
 DQ 571 GGQCVKYGDEGPKPTHGHSWSSWSPCSRTCCGGVSHRSLCTNPKPSHGKPCGSTR 630
 QY 668 TLKLCNSQKCPDSDVDFRAAQCAEHSNRFRGRHYKWKPYTOVEDQDLCKLYCIAEGFDF 727
 DQ 631 TLKLCNSQKCPDSDVDFRAAQCAEHSNRFRGRHYKWKPYTOVEDQDLCKLYCIAEGFDF 690
 QY 728 FFSLSNKVKGTGTPCSEDSNRVCIIDIGICERVGCNVLGSDAVEDVCGVCGNNSACTIHRG 787
 DQ 691 FFSLSNKVKGTGTPCSEDSNRVCIIDIGICERVGCNVLGSDAVEDVCGVCGNNSACTIHRG 750
 QY 788 LYTKHHTNQYHVMVTIPSGARSIRIYEMNVSTYSVRNALRYLYNGHWTVDWPGRYK 847
 DQ 751 LYTKHHTNQYHVMVTIPSGARSIRIYEMNVSTYSVRNALRYLYNGHWTVDWPGRYK 810
 QY 848 FSGTTFDYRRSNPEKNLIATGPTNETLIVELLPQGRNPGVAVESNPRIGTEKQPPAQ 907
 DQ 811 FSGTTFDYRRSNPEKNLIATGPTNETLIVELLPQGRNPGVAVESNPRIGTEKQPPAQ 870
 QY 908 SYTWAIVRSECVSCGGGR 926
 DQ 871 SYTWAIVRSECVSCGGGG 889

RESULT 2
 Q8K206 PRELIMINARY; PRT; 900 AA.
 AC Q8K206, 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC034739; AA34739.1; -
 DR InterPro; IPR002870; Rep_M12B_proprep.
 DR InterPro; IPR001590; RepPolysin.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF01562; Rep_M12B_proprep; 1.
 DR Pfam; PF01421; RepPolysin; 1.
 DR Pfam; PF00090; TSP1; 1.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PSS0215; ADAM_MEPRO; 1.
 DR PROSITE; PSS0092; TSP1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 900 AA; 100679 MW; 590BE2B0E73740FF CRC64;

Query Match 64.9%; Score 3400; DB 11; Length 900;
 Best Local Similarity 65.7%; Pred. No. 3.6e-294;
 Matches 646; Conservative 64; Mismatches 109; Indels 164; Gaps 13;

QY 9 RGLAALW-MLLAQVAVQVSGRSHQNRGSGOLEASPPRLLSRGRRLTAMSPLFSAGT 67
 DQ 4 RGLAALWMLLAQVAVQVSGRSHQNRGSGOLEASPPRLLSRGRRLTAMSPLFSAGT 67
 QY 68 CVRHGTSGGAWPERPASSSTRAGAGLDGKGRMDDEAGNHRSGQTNTGTENQTLHLVTQ 127
 DQ 46 -----SGSW-----LETG-----E 55
 QY 128 YDLVSAVEVDHGRDYVSHEIMHQRRAVAVSEV--ESLHLKLGPRHDFHMDLTSS 185
 DQ 56 YDLVSAVEVDHGRDYVSHEIMHQRRAVAVSEV--ESLHLKLGPRHDFHMDLTSS 115
 QY 186 LVAGGFTVOTLTKGTSVOTLPEDFCYQSLRSHRNSPSHGKCEGSTRILKLCNS 245
 DQ 116 LVAGGFTVOTLTKGTSVOTLPEDFCYQSLRSHRNSPSHGKCEGSTRILKLCNS 170
 QY 246 QKCPDSDVDFRAAQCAEHSNRFR-----GRHYKWKPYTOVEA-----DL 285
 DQ 171 -----TKDTYFLKPLPPLTSLKLNRSAGDSPSHVLYKRSTEPHAPGASEVLMTKRDL 236
 QY 286 CKLYCIAEGFDFPSLSNKVKGTGTPCSEDSNRVCIIDIGICELSVSTSAHNPQPKEDLFI 345
 DQ 227 ARPHLHNDHNLGSPQKHF-----CGRRKK-----YMPQPPDGLXI 264
 QY 346 LPDEYKSLRHRKSLRSHRNEELVETLVVVDKMKMNHGHNITTYVLTILNMYSAFP 405
 DQ 265 LPDEYKSLRHRKSLRSHRNEELVETLVVVDKMKMNHGHNITTYVLTILNMYSAFP 324
 QY 406 KD-----GLMKDGTGRHDAILL 423
 DQ 325 KDGTIGNINIVIGLILLEDEQGLAISHAADTLTFCQWQSLGKMGKGTGRHDAILL 384
 QY 424 TGLDICSWKNEPCDTLGFAPISGCKYRSCTINEDTGLGFTIAHESGHNFNMHDSGNMC 483
 DQ 385 TGLDICSWKNEPCDTLGFAPISGCKYRSCTINEDTGLGFTIAHESGHNFNMHDSGNMC 444
 QY 484 GNMCKSEGNIMSPFLAGNGVFSWSPCSROYLHKFLSTQAQICLADQPKVKEYKYPEK 543
 DQ 445 GNMCKSEGNIMSPFLAGNGVFSWSPCSROYLHKFLSTQAQICLADQPKVKEYKYPEK 504

QY 544 LPGLYDANTOCKWQFGEKAKLMLDFKDDICAKLMCHRIKRCETKFPAAEGTICGHD 603
 DB 505 LPGLYDANTOCKWQFGEKAKLMLDFKDDICAKLMCHRIKRCETKFPAAEGTICGHD 564
 QY 604 MWCRCGQCVKYGDEGPKPTTHGHSNDWSNSPCSTCGGVSUHSRLCTTPKPSHGKFCPE 663
 DB 565 MWCRCGQCVKYGDEGPKPTTHGHSNDWSNSPCSTCGGVSUHSRLCTTPKPSHGKFCPE 624
 QY 664 GSTTILKLSNOKCPDSVDFRAQCAEHSRRFRGRHYKWKPYTOVEQDQDLCKLYCIAE 723
 DB 625 GSTTILKLSNOKCPDSVDFRAQCAEHSRRFRGRHYKWKPYTOVEQDQDLCKLYCIAE 684
 QY 724 GDFDFFSLSNKVKDGTFCSDSRNVCIDGICERVCNDVLSGDAVEDVCGVNCNNSACT 783
 DB 685 GDFDFFSLSNKVKDGTFCSDSRNVCIDGICERVCNDVLSGDAVEDVCGVNCNNSACT 744
 QY 784 IHRGLYTKHHTNOYHHWVTIPSGARSIRIYEMNVSTYSISVSNALRRYYLNGHWTVDWP 843
 DB 745 THRLGYSKHSTNQYHHWVTIPSGARSIRIYEMNVSTYSISVSNALRRYYLNGHWTVDWP 804
 QY 844 GRYKFSGTTFYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAVESMPLRGTEKOP 903
 DB 805 GRYKFSGTTFYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAVESMPLRGTEKOP 864
 QY 904 PAQPSYTWAIVRSECSVSCGGGR 926
 DB 865 AAQPSYTWAIVRSECSVSCGGGR 887

RESULT 3

Q8TE60 PRELIMINARY; PRT; 1021 AA.
 AC Q8TE60;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein KIAA2029 (Fragment).
 GN KIAA2029.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CO NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB095949; BAC23125.1; -.
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 1021 AA; 113954 MW; 57EFSB95180C14AE CRC64;

Query Match 62.6%; Score 3280; DB 4; Length 1021;
 Best Local Similarity 92.9%; Pred. No. 2.2e-283;
 Matches 590; Conservative 2; Mismatches 1; Indels 42; Gaps 1;

QY 334 HWPQPPKEDFILPDEYKSLRHKRSLLSRSHNEELNVETLVVVDKQKQNGHENITY 393
 DB 52 YNPQPPKEDFILPDEYKSLRHKRSLLSRSHNEELNVETLVVVDKQKQNGHENITY 111
 QY 334 VLTILNMYSAFPKQ-----GLMG 411
 DB 112 VLTILNMYSAFPKQDTIGGINIAIVGLLLEDPGLVISHADHTLSFCQWQGLMG 171
 QY 412 KGTGTHDAIILTLGLDLSWNEPCDTLGFAPISGCMYSKYSTINEDTGLGFAFTIHE 471
 DB 172 KGTGTHDAIILTLGLDLSWNEPCDTLGFAPISGCMYSKYSTINEDTGLGFAFTIHE 231
 QY 472 SGHNFEMIHDGEGNMCKSEGNIMSPTLAGRNQVFSWSPCSRYLHKFLSTAQAICLADQ 531
 DB 232 SGHNFEMIHDGEGNMCKSEGNIMSPTLAGRNQVFSWSPCSRYLHKFLSTAQAICLADQ 291

QY 532 PKPVKEYYPEKLPGELYDANTOCKWQFGEKAKLMLDFKDDICAKLMCHRIKRCETKFP 591
 DB 292 PKPVKEYYPEKLPGELYDANTOCKWQFGEKAKLMLDFKDDICAKLMCHRIKRCETKFP 351
 QY 592 MPAAEGTICGHDWNCRCGQCVKYGDEGPKPTTHGHSNDWSNSPCSTCGGVSUHSRLCT 651
 DB 352 MPAAEGTICGHDWNCRCGQCVKYGDEGPKPTTHGHSNDWSNSPCSTCGGVSUHSRLCT 411
 QY 652 NPKPSHGKFCGEGSTRTLLKLSNOKCPDSVDFRAQCAEHSRRFRGRHYKWKPYTOVE 711
 DB 412 NPKPSHGKFCGEGSTRTLLKLSNOKCPDSVDFRAQCAEHSRRFRGRHYKWKPYTOVE 471
 QY 712 DQDLCKLYCIAEGDFDFFSLSNKVKDGTFCSDSRNVCIDGICERVCNDVLSGDAVEDV 771
 DB 472 DQDLCKLYCIAEGDFDFFSLSNKVKDGTFCSDSRNVCIDGICERVCNDVLSGDAVEDV 531
 QY 772 CGVNCNNSACTIHRGLYTKHHTNOYHHWVTIPSGARSIRIYEMNVSTYSISVSNALRR 831
 DB 532 CGVNCNNSACTIHRGLYTKHHTNOYHHWVTIPSGARSIRIYEMNVSTYSISVSNALRR 591
 QY 832 YYLNGHWTVDWPGRYKFSGTTFYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAV 891
 DB 592 YYLNGHWTVDWPGRYKFSGTTFYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAV 651
 QY 892 YSMPLRGTEKOPPAQPSYTWAIVRSECSVSCGGGR 926
 DB 652 YSMPLRGTEKOPPAQPSYTWAIVRSECSVSCGGGR 686

RESULT 4

Q8TE60 PRELIMINARY; PRT; 1081 AA.
 AC Q8TE60;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE ADAMTS18 protein.
 GN ADAMTS18.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CO NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Ohaya A.J., Llamazares M., Garabaya C., Quesada V.,
 Lopez-Otin C.;
 RT "Cloning, expression analysis, and structural characterization of
 seven novel human ADAMTSs, a family of metalloproteinases with
 disintegrin and thrombospondin-1 domains."
 RL Gene 483:49-62(2002).
 DR EMBL; AJ311903; CAC83612.1; -.
 DR Genbank; HGNC:17110; ADAMTS18.
 DR InterPro; IPR002870; Rep M12B prop. 1.
 DR InterPro; IPR001590; Repolysin.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF01562; Rep M12B prop. 1.
 DR Pfam; PF01421; Repolysin; 1.
 DR Pfam; PF00090; TSP 1; 4.
 DR SMART; SM00209; TSP1; 4.
 DR PROSITE; PS50215; ADAM MEPR0; 1.
 DR PROSITE; PS50092; TSP1; 3.
 SQ SEQUENCE 1081 AA; 119655 MW; 0438BF645676461E CRC64;

Query Match 46.0%; Score 2408; DB 4; Length 1081;
 Best Local Similarity 51.0%; Pred. No. 1.7e-205;
 Matches 463; Conservative 118; Mismatches 197; Indels 130; Gaps 18;

QY 85 ASSSTRGAAGLOGKGRDMEAGNHRSQNTGTGTENQTLVLTQYDLVSAYEVDHGRGDYVS 144
 DB 47 ASDSSSGASGLN-----DDYVFTPEVDSAGSYIS 77
 QY 145 HEIMHQRRRRANAVSEVSLHLRLKPRHDFMDLRTSSSLVAPGFIVQTLGKTGTSV 204


```

Db 78 HDILHNGRKKK-AQARSSSHAYRFAQGLHULK-PSAILSSHIVQVLKGDGASSET 135
Qy 205 QTLPPEDFCFYQGLSRSHRNPSHGKFCGEGSTRLTKLNSQ-----KCPRDSVDFFRAAQC 260
Db 136 OK-PEVOQCFYQGFTRND-SSSSAVASTCAGLSGLIRKNEFLISLPQ-----LLA 186
Qy 261 AHNSSRRRGRHYKWKPYTOVEADLCKLYCIAEGDFEFSSLNKVKDG-----TPCESDS 315
Db 187 QBNHSSPAG-HHPHVLRYRTAEKIQRYGPG-----SGRNYGYPSPHPIHASQS 238
Qy 316 R-----NVCIDGICELSVSTSAHMPQPKEDLFLPDEYKSLRHKRSLRS 363
Db 239 RETEYHRRLOKHFC-----GREKKYAPKPTEDTYLRDEYSGSRPRRSAGKS 289
Qy 364 HNEBELNVETLVVDKMMQWGHENITTVLTILNMVSALFKDG----- 408
Db 290 QKG--LNVETLVVADKMKVKEHGKGNVTYITLVK-VSLFKDGTIGSDINVVVSLIL 346
Qy 409 -----LMGKDGTRHDHAILLTGLDICSWKNEPCDTLGF 441
Db 347 LEQEGGLLNNHADQSLNSFCWQSQSALIGKNGKRHDHAILLTGFDICSWKNRPCDTLGF 406
Qy 442 APISCMSKYRSCNTNEDTGLAFTIAHESGHNFGMHGEGNMCKSEGNIMSPFLAG 501
Db 407 APTSCMSKYRSCNTNEDTGLAFTIAHESGHNFGMHGEGNCPCKRAEGNIMSPFLAG 466
Qy 502 RNVFWSWSPCSROYLHKFLSTAQAI CLADQPKPKYKPEKLPGLDYDANTCKWOFGE 561
Db 467 NNVFWSWSPCSROYLHKFLSTAQAI CLADQPKPKYKPEKLPGLDYDANTCKWOFGE 526
Qy 562 KAKLCHLDFKDI CKALWCHRIKRCETKMPAAEGTICGHDWMCRCGQCVKYGDGPKP 621
Db 527 KAKLCSLGFVKDICKSLWCHRVGHCETKMPAAEGTICGHDWMCRCGQCVKYGDGPKP 586
Qy 622 THGHWDSWSSPSCSRTCCGGVSHRSRLCTNPKPSHGKFCGEGSTRLTKLNSOKCPRDS 681
Db 587 IHGQWSANSKWSSECSRTCCGGVSHRSRLCTNPKPSHGKFCGEGSTRLTKLNSOKCPRDS 646
Qy 682 VDFRAAOCAEHSNRPRGRHYKWKPYTOVEDODLCKLYCIAEGDFEFSSLNKVKDTPC 741
Db 647 LDFRAAOCAEHSNRPRGRHYKWKPYTOVEDODLCKLYCIAEGDFEFSSLNKVKDTPC 706
Qy 742 SEDSRNVICDTCERVGCNVLGSDADEVCGVCGNNSACTIHRGLYTKHHHTNQYHYM 801
Db 707 SPKNQDVCLGVCELVGCHHELGSKAVSDACGVCKGDNSTCKFYKGLYLQHKANEYYPV 766
Qy 802 VTIPSGARSIRIYEMNVSTSVISURNALRRYVILNGHWTVDMPGRYKFSGTTFDYRSYNE 861
Db 767 VLIPAGARSIRIYEMNVSTSVISURNALRRYVILNGHWTVDMPGRYKFSGTTFDYRSYNE 826
Qy 862 PENLIATGPTNETLIVELLFOQNPFGVAWEYSMPRL--GTEKQPPA--QPSYTWAIVRSE 917
Db 827 PERLYAPGPTNETLIVELLFOQNPFGVAWEYSMPRL--GTEKQPPA--QPSYTWAIVRSE 883
Qy 918 CSVSCGGG 925
Db 884 CSVSCGGG 891

RESULT 5
ID Q8BKAL PRELIMINARY; PRT: 1092 AA.
AC Q8BKAL;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE Hypothetical metallopeptidase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN NCBI_TaxId=10090;
[1]

```

```

RP SEQUENCE FROM N.A.
RX STRAIN=CS7BL/6J; TISSUE=Eye;
RA MEDLINE=22354683; PubMed=12468851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK053851; BAC35556.1; --
KW Hypothetical protein.
SQ SEQUENCE 1092 AA; 120324 MW; BSF039565S3536AE CRC64;

Query Match 43.1%; Score 2257; DB 11; Length 1092;
Best Local Similarity 59.4%; Pred. No. 5.1e+19;
Matches 397; Conservative 90; Mismatches 117; Indels 64; Gaps 5;

Qy 313 EDSRNVICDTCERVGCNVLGSDADEVCGVCGNNSACTIHRGLYTKHHHTNQYHYMVTIPSGARS 372
Db 82 DDARNV-----YAPKPAEDAYLAFDEYGGTGRPRRSAGKSNG--LNVE 124
Qy 373 TLVVVDKMMQWGHENITTVLTILNMVSALFKDG----- 408
Db 135 TLVVADAKMKVKEHGKGNVTYITLVNMVSSLFKDGTIGSDINVVVSLILLEPEEGLL 184
Qy 409 -----LMGKDGTRHDHAILLTGLDICSWKNEPCDTLGFAPISGCMCK 450
Db 185 INNHADQSLNSFCWQSQSALIGKNGKRHDHAILLTGFDICSWKNRPCDTLGFAPISGCMCK 244
Qy 451 YPSCNTNEDTGLAFTIAHESGHNFGMHGEGNMCKSEGNIMSPFLAGRNGVFSWSP 510
Db 245 YPSCNTNEDTGLAFTIAHESGHNFGMHGEGNCPCKRAEGNIMSPFLAGRNGVFSWSS 304
Qy 511 CSROYLHKFLSTAQAI CLADQPKPKYKPEKLPGLDYDANTCKWOFGEKAKLCLMDF 570
Db 305 CSROYLHKFLSTAQAI CLADQPKPKYKPEKLPGLDYDANTCKWOFGEKAKLCLSGV 364
Qy 571 KKDICKALWCHRIKRCETKMPAAEGTICGHDWMCRCGQCVKYGDGPKPTHGHWSDWS 630
Db 365 MKDICKALWCHRIKRCETKMPAAEGTICGHDWMCRCGQCVKYGDGPKPTHGHWSDWS 424
Qy 631 SWPSCRTCCGGVSHRSRLCTNPKPSHGKFCGEGSTRLTKLNSOKCPRDSVDFFRAAQA 690
Db 425 KMSECSRTCCGGVSHRSRLCTNPKPSHGKFCGEGSTRLTKLNSOKCPRDSVDFFRAAQA 484
Qy 691 EHNSRRRGRHYKWKPYTOVEDODLCKLYCIAEGDFEFSSLNKVKDTPCSEDSRNVCI 750
Db 485 EYNNKPPFGWLYRWKPYTKVEEDRCKLYCKAENFEFFFANSKVKDGTFCSPHNDVCI 544
Qy 751 DGICERVGCNVLGSDADEVCGVCGNNSACTIHRGLYTKHHHTNQYHYMVTIPSGARS 810
Db 545 DGICELVGCNHELGSKAVSDACGVCKGDNSTCKFYKGLYLQHKANEYYPVTIPAGARS 604
Qy 811 IRIYEMNVSTSVISURNALRRYVILNGHWTVDMPGRYKFSGTTFDYRSYNEPENLIATGP 870
Db 605 IEIQELQSSSYLAVRSLSKYILTGWSIDWPGDFTFAGTTFEYQSRNRPRLATGP 664
Qy 871 TNETLIVELLFOQNPFGVAWEYSMPRLGTEKQPPAQP--SYTWAIVRSECSVSCGGGCLP 929
Db 665 TNETLIVELLFOQNPFGVAWEYSMPRLGTEKQPPAQP--SYTWAIVRSECSVSCGGGCLP 929
Qy 930 VLLLEAAC 937
Db 721 YISKAIC 728

RESULT 6
ID Q8BKX1 PRELIMINARY; PRT: 1009 AA.
AC Q8BKX1;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE ADAMTS-12 precursor (Fragment).

```


QY 898 GTEKQPPAQSPTWAIVR-SECVSCGGRCGLPVLILLERACOPSATA 943
 DB 786 ARDALPP-YSWHYAPWTKSAQAGSGSQVQVECRNQLDSSAVA 828

RESULT 8
 ID Q9W493 PRELIMINARY; PRT; 1059 AA.
 AC Q9W493;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE CG4096 protein.
 GN CG4096.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke Z., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodeon K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Feriz C., Ferreira S., Fleischmann W.,
 RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman J.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Munkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle B.J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Stadler A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu O., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL SEQUENCE 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,

PAcleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.P., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Beran B., Carlson J.W., Celniker S.E.,
 RA Clamp W., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003435; AAF46065.2; -;
 DR HSP; P15167; IATL;
 DR MEROPS; M12.231; -;
 DR FlyBase; FBGN0029791; CG4096.
 DR InterPro; IPR002870; Pep M12B_propep.
 DR InterPro; IPR001590; Repolysin.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR006025; Zn_WTpeptdase.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Repolysin; 1.
 DR Pfam; PF00090; TSP1; 2.
 DR SMART; SM00209; TSP1; 2.
 DR PROSITE; PS50092; TSP1; 2.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 SQ SEQUENCE 1059 AA; 119195 MW; A03C99E47618F3D7 CRC64;

Query Match 20.5%; Score 1075; DB 5; Length 1059;
 Best Local Similarity 28.3%; Pred. No. 1.6e-86;
 Matches 273; Conservative 132; Mismatches 309; Indels 250; Gaps 27;
 QY 138 HRCGYVSHETMHQRRRAVAVSEVES-----LHLRLKGRPHDFMDLRTSSSLVARGFTV 193
 DB 63 HQLEY-AHELDHRRHRRQRLNSLSEHDTQADLHLPLANETLHJELMAHSYFLAPNLVV 121
 QY 194 Q-----TLGKT 199
 DB 122 ERHRDLRTSLPTRLNCHFHGKVGQPATVVAISTCAGLVSCPNFCVLPELITLCQV 181
 QY 200 GTKSVQTLPEDF-----CFYQGSRL-----SHRNSPSHGKFCES 236
 DB 182 G-HIRTAGNEYFIEPSKEHEPHVNGHPVQSRSSVKPKSLRKNRKKGKGR--SGS 237
 QY 237 TRLKLCNSQKCRSDSVDFRAAQA-----EHSNRRFRGRHYKWK- 276
 DB 238 GAEVSNCGTRE-PRRRMETLEWQARKVKVQSGRQIRRRHHHHHHHHKVKYRHQKTI 296
 QY 277 ---PYTQVEADLCKLYCIAAGDFPFSLSNKVKDGTCPCEDSRNVCIDICELSVVSTSA 333
 DB 297 SRVPTTKFK----- 305
 QY 334 HMPQPPKEDLFLPDEVKSLRHKRLSRHNEELNVELTVVVDKMMQNHGHNITTY 393
 DB 306 -----YETQFQEPHAEIPRRRSI-----SSPRHVELIVADATM--SAFHDLNLCY 352
 QY 394 VLTILNVSALFQGLMG----- 411
 DB 353 LLTIMNMVSALYKDPISGNSIEIVVVRITIQLOEEESQLQLNLTQNAQKNLDRFCWSQHL 412


```

QY 412 -----KGTGTHDAHLLTGLDICSWKNEPCDTLGPAPISGCMCKSVRSTCTNEDTGLGAF 466
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS50215; ADAM_MEPPO; 1.
DQ 413 NKGSEKDPHHHVALITKMLCA---NNCWTLGLANVGCMCKPKQSCSNVEDNGIMLSH 469
DR PROSITE; PS50092; TSP1; 5.
QY 467 TIAHSGHNFMIHDEGCMCKKSG---NIMSPTLAQRNVFVSWPSCSYOYLHKFLSTA 523
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 470 TITHLGHNFHMGHDTAKTIGCHPRVGPVIMHTPTFGADTLQVCWNSCSRKYITHFDQOG 529
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
QY 524 QALCLADQPKVKEKYKPEKLGELYDANTCKWQFG--EKAKLMLDFKDKICKALWCH 581
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 530 LSECLDDPTPLDEINVTGELFOMRYNARGCQRLQFNLTTSSEVCAACSAPEFESTLNC- 588
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
QY 582 RIGRKCETKMPAAEGTICGHDMWCRGGOCVKYGDGPKPTHGHSWSSWSPSCRTCGG 641
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 589 KVNQGEVTHMRPTAFOTGLGRNKCQNGKCVR--REELAAVNGGNGWSESECSRCGG 646
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
QY 642 GVSHRSLCTNPKPSHGKFCGSTRTLKLNCSQKCPDSDVDFRAAOCAEHNSRRFRGRH 701
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 647 GVSTQRECDNVPVANGVFCIGERKRYKICRKPCEBPSFPAQCCARFDNVSQAT 706
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
QY 702 YKWKPYTOVEDDLCKLYCIAEGDFPFSLSNKVKDGTGTPCSEDSERNVCIDGICERVCDN 761
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 707 YKWLFP--PDKNPKCLFSDVDDTIANWGATVLDGTPCTLGTNNWCIDGICKVGCWD 764
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
QY 762 VLGSDAVEDVCGVCGNNSACTIHRGLYTKHHHTNQ--YHMTVTPSGARSIRIYEMNVS 819
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 765 IVQSEVQDRCGCGSGGQCPVRETYDPPAAGDGYVELVTIPARARHILIRELANS 824
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
QY 820 TSYISVR--NALRRYLLNGHMTVDWPGRYKFSGTTTFYRRSYNPFENLIATGPTNETLIV 877
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 825 PHFLATGCGDGRPYLNGDSLISMFGFEIAGASLYDR-VDEQITITIPQIHSISL 883
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
QY 878 ELIFQGR--NPGVAVESYMRPLGTQKPPAQSYTALIVR-SECSVSCGGRCCLPVILLE 934
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DQ 884 YAIRVNESNAGIFTEFLPALNV---TAGRQFQRLSNWTACSACGG---VOHRE 935
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
QY 935 AACQ 938
DQ 936 PICQ 939

```

RESULT 9

```

Q8TE59 PRELIMINARY; PRT; 1207 AA.
ID Q8TE59
AC Q8TE59
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE ADAMTS-19.
GN ADAMTS19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S.; Obaya A.J.; Llamazares M.; Garabaya C.; Quesada V.;
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTS, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains."
RL Gene 283:49-62(2002).
DR EMBL; AJ311904; CAC84565.1;
DR GenBank; U011904; ADAMTS19.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR006025; Znf_Mtpeptdse.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.

```

```

DR Pfam; PF00090; tsp_1; 5.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS50215; ADAM_MEPPO; 1.
DR PROSITE; PS50092; TSP1; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1207 AA; 134061 MW; AP36F6BF5086FDE2 CRC64;

Query Match          19.5%; Score 1018.5; DB 4; Length 1207;
Best Local Similarity 31.7%; Pred. No. 2.1e-81;
Matches 270; Conservative 110; Mismatches 296; Indels 175; Gaps 32;

QY 167 LRLKQPRHDFHMDLSTSSUVAPOFIYOTL-----GKTGTSYV--QTLPEDFCFYQCSL 219
DQ 167 LRLKQPRHDFHMDLSTSSUVAPOFIYOTL-----GKTGTSYV--QTLPEDFCFYQCSL 219
QY 220 RSHRNSPSH-----GGK-----FCGSTRTLKLNCSQKCPDSDVDFRAAOCAEH 263
DQ 220 RSHRNSPSH-----GGK-----FCGSTRTLKLNCSQKCPDSDVDFRAAOCAEH 263
QY 227 LSHPSGLASFSTCGGSLMGFIQLNEDFIIFPLNDTMAITG-----H 268
DQ 227 LSHPSGLASFSTCGGSLMGFIQLNEDFIIFPLNDTMAITG-----H 268
QY 264 NERRPRGRHYKWKPYQVQVADLCKLYCIAEGDFPFSLSNKVKDGTGTPCSEDSERNVCIDGI 323
DQ 264 NERRPRGRHYKWKPYQVQVADLCKLYCIAEGDFPFSLSNKVKDGTGTPCSEDSERNVCIDGI 323
QY 269 PHRVYRQKRSMBEKVTEKSA-LHSHYC---GI-----ISDK---GRP---RSRKIAESG- 312
DQ 269 PHRVYRQKRSMBEKVTEKSA-LHSHYC---GI-----ISDK---GRP---RSRKIAESG- 312
QY 324 CELSVSTSAHWPQPPKEDLFIPLDQYKSLRHRKSLRSLRSHRNEELNVEVLVVDKQMW 383
DQ 324 CELSVSTSAHWPQPPKEDLFIPLDQYKSLRHRKSLRSLRSHRNEELNVEVLVVDKQMW 383
QY 313 -----RGKYSYKLPQY-----NIEVVVADPMVS 339
DQ 313 -----RGKYSYKLPQY-----NIEVVVADPMVS 339
QY 384 NHGHENITTYVLTILNMVSAFLKQGLM----- 410
DQ 384 NHGHENITTYVLTILNMVSAFLKQGLM----- 410
QY 340 YHGADAARFILTILNMVFNLFQKSLGVQVNLRLVILILHETPELYTCHHGKML 399
DQ 340 YHGADAARFILTILNMVFNLFQKSLGVQVNLRLVILILHETPELYTCHHGKML 399
QY 411 -----GKGTGTHDAHLLTGLDICSWKNEPCDTLGPAPISGCMCKSVRSTCTNEDTGLGAF 447
DQ 411 -----GKGTGTHDAHLLTGLDICSWKNEPCDTLGPAPISGCMCKSVRSTCTNEDTGLGAF 447
QY 400 FCKNGHREFGKNDIHLKSNWGEDMTSDVAALITKDFCVKHDPCDTVGLIAYLSGM 459
DQ 400 FCKNGHREFGKNDIHLKSNWGEDMTSDVAALITKDFCVKHDPCDTVGLIAYLSGM 459
QY 448 CSKYRSTINEDTGLGAFTHAESHGNNFMHDEGCMCKKSG-NIMS-PTIAGRN-G 504
DQ 448 CSKYRSTINEDTGLGAFTHAESHGNNFMHDEGCMCKKSG-NIMS-PTIAGRN-G 504
QY 460 CSEKRKCTIAEDNGLNLAFTIAHBMHGNHNDHPSC--ADGLHMSGEMIKGNL 517
DQ 460 CSEKRKCTIAEDNGLNLAFTIAHBMHGNHNDHPSC--ADGLHMSGEMIKGNL 517
QY 505 VFSWSPCSROYLHKFLSTAQALCLAD-QPKPVKEKYKPEKLGELYDANTCKWQFGEKA 563
DQ 505 VFSWSPCSROYLHKFLSTAQALCLAD-QPKPVKEKYKPEKLGELYDANTCKWQFGEKA 563
QY 518 DVSNRCKEDLERFLRSKASNCNLLQTPQSVNSWVPSKLPFGMTYTTADQCQILFGLA 577
DQ 518 DVSNRCKEDLERFLRSKASNCNLLQTPQSVNSWVPSKLPFGMTYTTADQCQILFGLA 577
QY 564 KLCLMLDFKDKICKALWCHRIGRK-CETKMPAAEGTICGHDMWCRGGOCVKYGDGPKPT 622
DQ 564 KLCLMLDFKDKICKALWCHRIGRK-CETKMPAAEGTICGHDMWCRGGOCVKYGDGPKPT 622
QY 578 SFCQ-EMOHVICTGLWCKVGEKECTKLDPPMDGTDCLGKWKCKAGECTS-RTSAPEHL 635
DQ 578 SFCQ-EMOHVICTGLWCKVGEKECTKLDPPMDGTDCLGKWKCKAGECTS-RTSAPEHL 635
QY 623 HGHWSDDSWSPCSRTCGGYSVSHRSLCTNPKPSHGKFCGSTRTLKLNCSQKCPDSDV 682
DQ 623 HGHWSDDSWSPCSRTCGGYSVSHRSLCTNPKPSHGKFCGSTRTLKLNCSQKCPDSDV 682
QY 636 AG---EWLWSPCSRTCSAGISSREKC--PGLDSEARDGNGPRKQYRICENPFCPAGLP 690
DQ 636 AG---EWLWSPCSRTCSAGISSREKC--PGLDSEARDGNGPRKQYRICENPFCPAGLP 690
QY 683 DFLAQAQCAEHNSRRFRGRH-YKWKPYTOVEDDILCKLYCIAEGDFPFSLSNKVKDGTGTPC 741
DQ 683 DFLAQAQCAEHNSRRFRGRH-YKWKPYTOVEDDILCKLYCIAEGDFPFSLSNKVKDGTGTPC 741
QY 691 GFRDMQCOAYSVRTSSPKHILQWQ--AVLDEKFCALFCSVPVGEQPIILSEKVMGDTSC 748
DQ 691 GFRDMQCOAYSVRTSSPKHILQWQ--AVLDEKFCALFCSVPVGEQPIILSEKVMGDTSC 748
QY 742 SEDSRNVCIDGICERVGCDNVLGSDAEDVCGVCGNNSACTIHRGLYTKHHHTNQYH 801
DQ 742 SEDSRNVCIDGICERVGCDNVLGSDAEDVCGVCGNNSACTIHRGLYTKHHHTNQYH 801
QY 749 GYQGLDICANGRCQKVGCDGLGLSLAREDHGCVNGKCKSKIIKGF--NHTRGAGVVEV 807
DQ 749 GYQGLDICANGRCQKVGCDGLGLSLAREDHGCVNGKCKSKIIKGF--NHTRGAGVVEV 807
QY 802 VTIPSGARSIRIYEMNVSSTYSVNRALRYLLNGHMTVDWPGRYKFSGTTTFYRRSYNE 861
DQ 802 VTIPSGARSIRIYEMNVSSTYSVNRALRYLLNGHMTVDWPGRYKFSGTTTFYRRSYNE 861
QY 808 LVIPAGARRIKVVEEKPAHSYLALRDA-GKQINSNDWKIEHSGAFNLAGTIVHYVRR-GL 865
DQ 808 LVIPAGARRIKVVEEKPAHSYLALRDA-GKQINSNDWKIEHSGAFNLAGTIVHYVRR-GL 865
QY 862 PENLIATGPTNETL-IVELAFQGNPVGVAWEYS-----MPRLCTEKOPPAQSYTALIVR 915
DQ 862 PENLIATGPTNETL-IVELAFQGNPVGVAWEYS-----MPRLCTEKOPPAQSYTALIVR 915
QY 866 WEKISAKGPTAPLHLLVLLFQDQYGLHYETIIPSDPLPENQSSKAP--EPLFWMTHTS 923
DQ 866 WEKISAKGPTAPLHLLVLLFQDQYGLHYETIIPSDPLPENQSSKAP--EPLFWMTHTS 923
QY 916 -SECSVSCGG 925
DQ 916 -SECSVSCGG 925
QY 924 WEDCDATCGG 934
DQ 924 WEDCDATCGG 934

```



```

RESULT 10
Q8NE26 PRELIMINARY; PRT; 967 AA.
AC Q8NE26;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC036515; AAH36515.1; -.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR006025; Zn_M12B_propep.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; csp_1; 2.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSPI; 3.
DR PROSITE; PS0215; ADAM_MEPPO; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR PROSITE; PS00092; TSPI; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 967 AA; 103387 MW; FFID399674201C3D CRC64;

Query Match 18.8%; Score 984.5; DB 4; Length 967;
Best Local Similarity 30.9%; Pred. No. 1.6e-78;
Matches 270; Conservative 104; Mismatches 328; Indels 173; Gaps 35;

QY 157 VAVSEVE-----SLHLRLKGRPHDFHMDLRTSSSLVAPGFIQTILG-KTGKSVQTLPP 209
DB 60 LVVLELAPGHGTRTLRLHAFDQDLDELPPSSFLAPGFTLQNVGRKSGSET--PLPE 117
QY 210 EDF--CFYQGLRHRNPSHGKFCGSGSTRTLKLSN-----QKCPDSVDFRAACAEH 263
DB 118 TDLACFYSGTVNGDPSSAA--ALSICEGVRCFAYLLGEAYFIQPLPAASRLATAAPGE- 175
QY 264 NSRRFRGRHYKWKPVTOVEADLCKLYCIAEGFDFFLSLNKVKDGTGTPCSDSRNVICDI 323
DB 176 -----KPPAPLQ-----FHLRLNRQGD-----VGGT 197
QY 324 CELSVVSTSAH-----MPQPKEDUFIPLDEYKSLRHKRSL 360
DB 198 C--GVVDDEPRPTGKAETEDDEGTGEGDEGQSPQDPALQGVQPTGTGS-IRKKR-F 253
QY 361 LRSRNEELNVETLVVVDKMNQNHENITTVTLNVLVSALFK----- 406
DB 254 VSSHR-----VYETMLVADQSMABFHG-SGLKHYLLTLFSAARLYKHPSTRNVSLSLVVK 308
QY 407 -----DGLMG-----KDGTRHDAHLLTLGLDICSWKNEP 435
DB 309 ILVHDEQKGEVTSNAAALTNRNFWNQKQHNPPSDRDAEHYDTAILFTQDLG--SQT 366
QY 436 CDITGAFISGCMCKYRSCCTNEDTGLAFTIAHSGHNFMGHDEGNCKKSEG--- 492
DB 367 CDITGMAVGVTCVPSRSCSVIEDDGLQAFTTAHELGHVFNPHD--DAKQASLNGVNQ 425
QY 493 --NIMSPTLACRNGVFSKSCROYLHKFLSTAQAICLADQPKVKYKYPEKLPGLDYD 550
DB 426 DSHWASMLNLSHQPNPSPSCAYWITISFLDNGHGECLMD--KQNPFIQLPGDIFGTSYD 483
QY 551 ANTQCKWQFGEKAKLCMLDFKKDICKALWCHRIGRK-----CETKFWPAAEAGTICGHDW 605

```

```

DB 484 ANRQOFTFGEDSKHC--PDAASTCSTLMC--TGTSGGVLCVCTKGFPPWADGTSCGEGKW 539
QY 606 CRGGQCVKYGDGEG--PKPTHGHSNDSSWSPSCRTCGGVSHRSRLCTNPKPSHGKAFCE 663
DB 540 CINGCNVKTDRKHFDTPFHGSGWMPGWDGCSRTGGGVQYTRCENDPVPKNGKYCE 599
QY 664 GSTRTLKLSNQKCP-RDSVDFRAACAEHN-----SRRFRGRHYKWKP-YTQVEDDQDLCK 717
DB 600 GKRVRYSNCLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCK 659
QY 718 LYCTAEGFDFPFLSNKVKGTGCTSDSRNVICDGCICERVGCNVLGSDADEVDCVCGNG 777
DB 660 LICQAKGIGYFFVLQPKVWDGTPCSTDTSTSVCGQCVKAGCDRIIDSKKKFKCKGCGG 719
QY 778 NNSACTIHRGLYTKHHHTNQYHVMVTIPSGARSIRIYEMNV-----STSVISVNRNALRY 832
DB 720 NGSTCKKISGSVTS--AKPGYHDIITPTGATNIEVKQRNQRNNGSFLAICAADGTY 777
QY 833 YLNGHWTVDMPCR-YKFSGTTFYRRSYNEPENLIATGPTNETLIVELLEQGR--NPGVA 889
DB 778 ILNGDYTLSTLEQDINMYKGVVLYRSGSSAALERIRSFSLKPELTIQVLTVGNALRPKIK 837
QY 890 WEYSMPRLGTGKQPPAOPESYT-WAIVR-SECVSC 922
DB 838 YTFPVKK--KKESEFNAIPTFSANVIEEWGECSSKSC 870

RESULT 11
Q19791 PRELIMINARY; PRT; 2165 AA.
AC Q19791; Q27524;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadasy S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
RA Craxton M., Dear S., Du Z., Durbin R., Faveillo A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showken R.,
RA Smailon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinscock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38 (1994).
DR EMBL; Z69361; CAA93288.1; -.
DR EMBL; Z69360; CAA93288.1; JOINED.
DR EMBL; Z69360; CAA93287.1; -.
DR EMBL; Z69361; CAA93287.1; JOINED.
DR HSSP; P15167; IDTH.
DR MEROPS; M12.135; -.
DR WormPep; F25H8.3; CE05729.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR006025; Zn_M12B_propep.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; csp_1; 14.

```


[illegible]

RESULT 13

Q8BZD1
ID Q8BZD1 PRELIMINARY; PRT; 497 AA.
AC Q8BZD1
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Hypothetical metalloprotease.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK035797; BAC29190.1; -.
KW Hypothetical protein.
SQ SEQUENCE 497 AA; 54372 MW; 3A72EDFC3988A289 CRC64;

Query Match 16.7%; Score 873.5; DB 11; Length 497;
Best Local Similarity 39.2%; Pred. No. 4.9e-69;
Matches 213; Conservative 54; Mismatches 133; Indels 143; Gaps 18;

QY 68 CVRHGTRGSAWPE-----RPSSTGCAAGLCKGKGRMDGAGNH 108
DB 9 CALRAAGGPPWAGLGRUAKALQLCCFCAVAVALASDSSG--GSGLND----- 61
QY 109 RSQONTGTENQTLHLVLTQDLYSAVEVDRHGVDYSHIMHQRRAVAVAVEVSLHLR 168
DB 62 -----YFVVVPEVDSGSGYISHDILHHRKRSAHAGS--NSLHYR 100
QY 169 LKGRPHDFMDLRTSSSLVAPGFTVOTLGKTKSVQTLPE-DFCFYQGSLSRSHNSPS 227
DB 101 VSARQDLHLEK-PSAILSSHRVQLGKDGAS--ETREPEVQCLYQGFTRND-SSSS 156
QY 228 HGGKFCGSTRTLKLCNSQKCPDSVDVFRAAQ-----CAEHNSRRFRGRHYKWKPYTOVE 282
DB 157 VAVSTCAGLSGLIR-----TRDN-EFLISPLPQLLAQEHNVSSPAG-HPHVLYKRTA 207
QY 283 ADLCKLYCIAGCPDFPESLKNKYKG-----TPCSSEDS-----RNVCIDGICE 325
DB 208 EKRVRVQDYFG-----SORTVPGHSPHTPPASQSOEPEYSHRWKRRHFC----- 254
QY 326 LSVVSTSAHMPQPPKEDFLIPDEYKSLRHRSLRSHRNELNVETLVVVDKMMQNH 385
DB 255 --GRRKYAPKPAEDAYLRDEYGTGRPREAGKSGNQ--LNVETLVVADAKWVKH 309
QY 386 GHENITTYVLTILNMVSALEKDG-----GSGVSDVYVSVLSLSEEPGGLLNHHADQSLNSFC 408
DB 310 GKDDVTYVLTILNMVSSLEKDG-----GSGVSDVYVSVLSLSEEPGGLLNHHADQSLNSFC 369
QY 409 -----LMGKDGTRHDHALITGLDICSWKNEPCDTLGFAPISGMSKYRSCITNEDTGLG 463
DB 370 QWQSALYQNGKRDHALITGLDICSWKNEPCDTLGFAPISGMSKYRSCITNEDTGLG 429
QY 464 LAFTIAHESGHNFGHIDGNGMCKKSEGNIMSPTLAGNGVSWSCROYLHKPLSTA 523
DB 430 LAFTIAHESGHNFGHIDGNGMCKKSEGNIMSPTLAGNGVSWSCROYLHKPLSTA 489
QY 524 QAI 526
DB 490 RCI 492

RESULT 14

Q8K384

Q8K384 PRELIMINARY; PRT; 833 AA.
AC Q8K384
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Similar to a disintegrin-like and metalloprotease (repolysin type)
DE with thrombospondin type 1 motif, 4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027773; AAH27773.1; -.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR006025; Zn_MTPeptase.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp.1; 1.
DR SMART; SM00608; ACE; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PSS0215; ADAM_MEPPO; 1.
DR PROSITE; PSS0092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Integrin; Protease; Metalloprotease.
SQ SEQUENCE 833 AA; 90097 MW; 4CEC83DFBC3AA619 CRC64;

Query Match 16.7%; Score 872.5; DB 11; Length 833;
Best Local Similarity 34.4%; Pred. No. 1.3e-68;
Matches 216; Conservative 77; Mismatches 230; Indels 105; Gaps 23;

QY 371 VETLVVVDKRMQNHGHEITTYVLTILNMVSALEKDG-----GLMCKDGTGRDH-----AILTGLDICSWKNEPCDTLGPAPIS 445
DB 216 VETLVVADDKWAAPFHG-TGLKRYLLTVMMAAAKAFKHPISRNPNVLNVLVTRLVILGSGQEG 274
QY 408 -----GLMCKDGTGRDH-----AILTGLDICSWKNEPCDTLGPAPIS 445
DB 275 PQVGFSAQTLRFCTWQRLNTPNDSDDPHPTAILTRQDLGCVST-CUTLGNWADVG 332
QY 446 GMCKSVRSCITNEDTGLGLAFTIAHESGHNFGMIHD-----GEGNMCKKSEGNIMS 496
DB 333 TVCDPAPSCAIVEDDGLQSAFTAAHGLGVFNLMHDSKPCITNLNGSGG-----SSRHVA 388
QY 497 PTLAGNGVSWSCROYLHKPLSTAQAICLADQKPKYKEYKPKLPELYDANTOCK 556
DB 389 PVAHVDPPEFWSPCSARFIFDLDNGYGHCLLD--KPEAPLHLPLATFFGKDYADQRCQ 446
QY 557 WQFGEKAKLWLDKFKDICKALWC--HRIGR-KCETKPMFAAEGTTCGHDMWCRGQCQVK 613
DB 447 LTFGPDSSHC--POLPPCAALWCSGHLNCHAMCQTKHSPWADGTPCGSSQACWGRCLH 504
QY 614 YGD--EPPKPTHGSDWSWSPCSTTCGGVSHRSLCTNPKPSHGKFCGSGSTRTLKL 671
DB 505 VDQLKDFNVFQAGSGWGPWGCSCRTCCGGVQVFSRDCRTPVPRNGKYCEGRKTRFRS 564
QY 672 CNSQKCPEDS-VDFRAAQAQCAHNSR-----RFRGRHYKWK-PYTOVEDODLCKLYCIABG 724
DB 565 CNTENCCHGSAITFREEQCAAYNHTDLFKSPG-PMDVVPRTYGVAPRDQCKLTCQARA 623
QY 725 FDFPFLSNKVKDGTCPGSEDSRNVICDGCICERVCNVLGSDADEVDCVCGNNSACTI 784
DB 624 LGYVYVLEPRVADGTPCSPDTSVVCVQGRCTHAGCDRIIGSKKKKFKDKMVCVCGDGRCSK 683
QY 785 HRGLYTKHHTNOYHMTIPSGARSIRIYEM--NVSTSYLSVRNARYLVINGHWT-- 839
DB 584 QSGSKAFRY--GYSDVVTIPAGATHILVRQCGSGSKSIYIALKLSGGSYALNGSYITLM 741
QY 840 -----VDWPG--RYKTSGLTTFDYRRSYNEFENIATGPTNETLIVELLFQG--RNPGVAV 890
DB 742 PSPTDVVLPGAVSLRYSGAI-----AASETLSGHGFLAQPLTLQVLVAGNPQARLRY 794

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 23:39:29 ; Search time 53 Seconds
(without alignments)
759.999 Million cell updates/sec

Title: US-09-981-151A-8
Perfect score: 5236
Sequence: 1 MKPRARGWGLAALWMLLAQ.....LEAACQPSATYATLAFLES 952

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/6C COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3846.5	73.5	1224	US-09-930-872-4	Sequence 4, Appli
2	1543.5	29.5	491	US-09-930-872-2	Sequence 2, Appli
3	1188	22.7	859	US-09-369-364A-5	Sequence 5, Appli
4	1179	22.5	997	US-09-369-364A-7	Sequence 7, Appli
5	1057.5	20.2	1081	US-09-369-364A-17	Sequence 17, Appli
6	1056	20.2	1882	US-09-369-364A-13	Sequence 13, Appli
7	991	18.9	874	US-09-369-364A-15	Sequence 15, Appli
8	984.5	18.8	967	US-09-130-491-2	Sequence 2, Appli
9	958.5	18.3	727	US-09-445-023A-12	Sequence 12, Appli
10	934	17.8	727	US-09-445-023A-1	Sequence 1, Appli
11	933	17.8	1205	US-09-491-522-11	Sequence 11, Appli
12	923	17.6	1211	US-09-491-522-5	Sequence 5, Appli
13	903	17.2	608	US-09-130-491-13	Sequence 13, Appli
14	885.5	16.9	930	US-09-369-364A-2	Sequence 2, Appli
15	878	16.8	930	US-09-122-126B-15	Sequence 15, Appli
16	878	16.8	930	US-09-634-286A-2	Sequence 2, Appli
17	856.5	16.4	837	US-09-122-126B-2	Sequence 2, Appli
18	856.5	16.4	837	US-09-634-286A-9	Sequence 9, Appli
19	823.5	15.7	905	US-09-369-364A-2	Sequence 2, Appli
20	778.5	14.9	551	US-09-130-491-16	Sequence 16, Appli
21	675.5	12.9	518	US-09-369-364A-22	Sequence 22, Appli
22	607	11.6	481	US-08-130-491-8	Sequence 8, Appli
23	470	9.0	566	US-09-491-522-7	Sequence 7, Appli
24	451.5	8.6	525	US-09-369-364A-21	Sequence 21, Appli
25	323	6.2	450	US-09-369-364A-19	Sequence 19, Appli
26	270.5	5.2	616	US-09-608-790-1	Sequence 1, Appli
27	268.5	5.1	751	US-08-836-443-3	Sequence 3, Appli

28	245.5	4.7	802	4	US-09-632-098-2	Sequence 2, Appli
29	245.5	4.7	812	4	US-09-632-098-4	Sequence 4, Appli
30	244.5	4.7	814	3	US-09-813-819-4	Sequence 4, Appli
31	244.5	4.7	814	4	US-09-813-819-2	Sequence 2, Appli
32	232.5	4.4	855	3	US-09-813-819-2	Sequence 2, Appli
33	232.5	4.4	855	4	US-09-920-048-2	Sequence 19, Appli
34	232	4.4	1172	1	US-08-313-288B-19	Sequence 20, Appli
35	228	4.4	1170	1	US-08-313-288B-20	Sequence 2, Appli
36	217	4.1	748	3	US-08-920-234-2	Sequence 9, Appli
37	215	4.1	748	3	US-09-030-335-9	Sequence 4, Appli
38	215	4.1	769	1	US-08-243-542-4	Sequence 4, Appli
39	215	4.1	769	1	US-08-477-407-4	Sequence 4, Appli
40	215	4.1	769	2	US-08-484-355-4	Sequence 3, Appli
41	212.5	4.1	529	2	US-08-836-442-3	Sequence 11, Appli
42	208	4.0	245	4	US-09-369-364A-11	Sequence 3, Appli
43	205	3.9	441	3	US-08-985-526-3	Sequence 14, Appli
44	204.5	3.9	464	3	US-09-411-329C-14	Sequence 3, Appli
45	203	3.9	670	1	US-08-243-542-3	Sequence 3, Appli

US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930.872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match 73.5%; Score 3846.5; DB 4; Length 1224;
Best Local Similarity 75.4%; Pred. No. 0;
Matches 738; Conservative 25; Mismatches 73; Indels 143; Gaps 13;

Qy	1	MKPRARGWGLAALWMLLAQVAEQVSPGRSHQGRNRSQGLEASPPRLLSRGPRRLTAMS	60
Db	1	MKPRARGWGLAALWMLLAQVAEQAP	31
Qy	61	PLFSGAGTCVRHGRSGSAWEPRPASSSTRGAAGLDGKRDMDGAGNHRSQQTNTGTENQ	120
Db	32	FAAAA	58
Qy	121	TLHLVLTQDLSAYEVDRGDVVSHEIMHQRRAVAVSEVESHLLKGRPHDFHMDL	180
Db	59	-----EVDLSAYEVDRGDVVSHEIMHQRRAVAVSEVESHLLKGRPHDFHMDL	112
Qy	181	RTSSSLVAPGFVQTLGKTGKSVQTLPPEDPCFYQGSLSRHSRNSPSHGKCEGSTRTL	240
Db	113	RTSSSLVAPGFVQTLGKTGKSVQTLPPEDPCFYQGSLSRHSRNS-SVALSTCOGLSGMI	171
Qy	241	KUCNSQKCPDSDVDFRAACAEHNSRRPRGR	291
Db	172	R-----TEADYFLRPLPSHLSWKL-GRAAQSSPSHVLYKASTPHAFGASEVLVT	222
Qy	292	ASGDFDFFFS--LSNKKVKDGTPCSDESRNVCIDGICELSVSTSAHMPQPPKEDLFLPDE	349
Db	223	STWELAHQPLHSSDLRLGLP---QKQHFC-----GRKKYMPQPPKEDLFLPDE	270
Qy	350	YKSCSLRHKRSLLSRHNEELNVETLVVVDKKQMMQNHGHENITTYVTLILNMVVSALFKD---	407

ALIGNMENTS

RESULT 1

Db 271 YKSLRHKSLRSHRNEELVETLVVDKMMQNHGHEHNTTYVLTILNMVSAFLFKGT 330
QY 408 -----GLMGKDGTRHDHAILLTGLD 427
Db 331 IGGNINIAIVGLLEDEQPLVISHHADHTLSSFCQWQSLMGKDGTRHDHAILLTGLD 390
QY 428 ICSWKNPECDTGLFAPISGMCSKYRSCNTINEDTGLGLAFTIAHESGHNFMIHDGEGNMC 487
Db 391 ICSWKNPECDTGLFAPISGMCSKYRSCNTINEDTGLGLAFTIAHESGHNFMIHDGEGNMC 450
QY 488 KSEGINMSPPTLAGRNGVFSWSPCSQYHLKFLSTAQALCLADQPPVKEYKPEKLPG 547
Db 451 KSEGINMSPPTLAGRNGVFSWSPCSQYHLKFLSTAQALCLADQPPVKEYKPEKLPG 510
QY 548 LYDANTQCKWQFGEKAKLMLDKFKDICKALWCHIRGRKCEKTFMPAABGTICGHDMWCR 607
Db 511 LYDANTQCKWQFGEKAKLMLDKFKDICKALWCHIRGRKCEKTFMPAABGTICGHDMWCR 570
QY 608 GGCQVKYGDGPKPTHGHSWSDWSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 667
Db 571 GGCQVKYGDGPKPTHGHSWSDWSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 630
QY 668 TLKLNKSNQKCPDSDVDFRAAQCAEHNSRRFRGRHYKWKPYTQVEDQDLCKLYCIAEGPDE 727
Db 631 TLKLNKSNQKCPDSDVDFRAAQCAEHNSRRFRGRHYKWKPYTQVEDQDLCKLYCIAEGPDE 690
QY 728 FFLSLNKVKDGTCCSDSDSRNVCIDGICERVCNDVLGSDADEVCGVGNNSACTIHRG 787
Db 691 FFLSLNKVKDGTCCSDSDSRNVCIDGICERVCNDVLGSDADEVCGVGNNSACTIHRG 750
QY 788 LYTKHHTNQYHYMVTIPSGARSIRIYEMNVSTSYISVRNALRYYLNGHWTVDWPGRYK 847
Db 751 LYTKHHTNQYHYMVTIPSGARSIRIYEMNVSTSYISVRNALRYYLNGHWTVDWPGRYK 810
QY 848 FSGTTFDYRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAVWEYSNPRLGTEKQPPAQ 907
Db 811 FSGTTFDYRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAVWEYSNPRLGTEKQPPAQ 870
QY 908 SYTWAIVRSECVSCGGGR 926
Db 871 SYTWAIVRSECVSCGGGQ 889

RESULT 2
US-09-930-872-2
; Sequence 2, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 491
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-2

Query Match 29.5%; Score 1543.5; DB 4; Length 491;
Best Local Similarity 57.7%; Pred. No. 1e-129;
Matches 333; Conservative 26; Mismatches 75; Indels 143; Gaps 13;

QY 1 MKPRARGWGLAALWMLLAQVAEQVSPGRSHQRNGSGQLEASPPRLSRGPRRLTAMS 60
Db 1 MKPRARGWGLAALWMLLAQVAEQP-----ACAMG 31

QY 61 PLFSAGTCVRHGTSGSAAWEPERPASSSTRGAAGLDGKGROWDEAGNHSQQTNTGTENQ 120
Db 32 PAAAA-----PGSPSVPRPPPPAERPG-----WMEKG-----58
QY 121 TLHLVLTQYDLVSAYEVDHRGDYVSGHEIMHQRRRRAVAVSEVESLHLRLKGRPHDFHMDL 180
Db 59 -----EYDLVSAYEVDHRGDYVSGHEIMHQRRRRAVAVSEVESLHLRLKGRPHDFHMDL 112
QY 181 RTSSSLVAPGFIQVTLGKTGTSVOTLPPEDFCFYQGSLSRHSRNSPSHGKFCGSTR 240
Db 113 RTSSSLVAPGFIQVTLGKTGTSVOTLPPEDFCFYQGSLSRHSRNS-SVALSTCQGLSGMI 171
QY 241 KLCNSQKCPDSDVDFRAAQCAEHNSRRFRGR-----HYKWKPYTQVEA-DLCKLYCI 291
Db 172 R-----TEADYFLRPLPSHLSWKL-GRAAQSSPSHVLVYKSTBPHAFGASEVLVT 222
QY 292 AEGFPDFPS--LSNKVKDGTCCSDSDSRNVCIDGICELSVWSTSAHMPQPKEDLPILPOE 349
Db 223 SRTWELAHQPLHSSDLRLGLP---QKQHFC-----GRRKKYMPQPKEDLPILPOE 270
QY 350 YESCLPHKSLRSLRSHRNEELAVETLVVDKMMQNHGHEHNTTYVLTILNMVSAFLFKD-- 407
Db 271 YKSLRHKSLRSHRNEELAVETLVVDKMMQNHGHEHNTTYVLTILNMVSAFLFKGT 330
QY 408 -----GLMGKDGTRHDHAILLTGLD 427
Db 331 IGGNINIAIVGLLEDEQPLVISHHADHTLSSFCQWQSLMGKDGTRHDHAILLTGLD 390
QY 428 ICSWKNPECDTGLFAPISGMCSKYRSCNTINEDTGLGLAFTIAHESGHNFMIHDGEGNMC 487
Db 391 ICSWKNPECDTGLFAPISGMCSKYRSCNTINEDTGLGLAFTIAHESGHNFMIHDGEGNMC 450
QY 488 KSEGINMSPPTLAGRNGVFSWSPCSQYHLKFLSTAQ 524
Db 451 KSEGINMSPPTLAGRNGVFSWSPCSQYHLKFLSTAQ 487

RESULT 3
US-09-369-364A-5
; Sequence 5, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-6
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (450)
; OTHER INFORMATION: Xaa = L
US-09-369-364A-5

Query Match 22.7%; Score 1188; DB 4; Length 859;
Best Local Similarity 31.4%; Pred. No. 1.8e-97;
Matches 284; Conservative 132; Mismatches 289; Indels 200; Gaps 25;

QY 108 HRSQQTNTGTENQTHLVLTQYDLVSAYEVDHRGDYVSGHEIMH--HQRRAV-----AYSE 161
Db 22 HSDHLSYSSQBEFTLYLEHYQLTPIRVQDQAGFLSFTVRKNDKHSRRSRMDPQQA 81
QY 162 VESLHLRLKGRPHDFHMDLRTSSSLVAPGFIQVTLGKTGTSVQTLPE-----DFCFY 215
Db 82 VSKLFFKL SAYKXKHLNLTNTDFVSKHFTVEYWGKG-----PQMKHDFLDNCHY 133

QY 216 QGSLRSHRSPSHGKFCFEGSTRITLCLNSQKCPROSDVFRAAOCAE-HNSRRFRGRHYK 274
Db : : : : :
QY 134 TGYLQDQR-----STTKVALSN-----CVGLRGVIATEDDEEYF 166
Db : : : : :
QY 275 WPKYTOVEADICKLYCIAEGPFFSLSNKVKDGT-----CSDSRNVICIDIGICLSV 328
Db : : : : :
QY 167 IEPLKNTT-----SKHFSY-----ENGHPHYVYKKSALQOQHLDYHSHCGVSD 211
Db : : : : :
QY 329 VTSAMHPOPPKEDLFIIPDEYKSC---LRHKRSLRSHRNEELNVETLVVVDKMMQNH 385
Db : : : : :
QY 212 FTRSGKPMWNLDTPLFIHYQINNTIHHROKRSV-----SIEFVETLVVADKMMGYH 266
Db : : : : :
QY 386 CHENITTVTLNVMVSAFLKDLGK-----411
Db : : : : :
QY 267 GRKDIHYTLVMNIVAKLYRDSLSGNVNIIVARLIVLTEDQPNLEINHADKSLSPFC 326
Db : : : : :
QY 412 -----KDG-----PRDHAILLTGLDICSWKNEPCTDLPAPISGCMKYRS 453
Db : : : : :
QY 327 KWQKSTLGHQSDGNTIPENGIAHHDNAVILITRYDICTYKNKPCGTLGLASVAGMCEPERS 386
Db : : : : :
QY 454 CTINEDTGLAFTTAHESCHNFGMIHDEGNMCKKSEGNIMSPTLAGNNGVFSWSPCSR 513
Db : : : : :
QY 387 CSINEDIGLGSFTTAHETVHNFVNMHDIIGNSC-----GR-----422
Db : : : : :
QY 514 QYLHKFSLTAQAICLADQPKPKYKYPEKLPGLDYDANTQCKWQFGEKAKLCLM-----568
Db : : : : :
QY 423 -----KWKQ-----QNYGSHYCEYQ---SFFLVCLQSRXH 452
Db : : : : :
QY 569 DFKKDICKALMCHRIORCKETKMPAAEGTICG-----HDMWCRGGQCVKYDEGPKPTHG 624
Db : : : : :
QY 453 QLFREVCRLWCLSKSNRCVTNSIPAAEGTLCQTGNIKMGWCYQGDVPPG-TWPSQIDG 511
Db : : : : :
QY 625 HNSDMSWSPCSTCGGVSYSRSLCTNPKPSHGKFCFEGSTRITLCLNSQKCPROSDVDF 684
Db : : : : :
QY 512 GHGWSLWGECSRTCGGVSYSRSLCHDCSPAPSEVEXYCLGERKXRSCTWPCPLGSRDF 571
Db : : : : :
QY 685 RAAQCAHNSRFRGRHYKWKPYTQVEDODLCKLYCIAEGPFFSLSNKVKDGTFCPSD 744
Db : : : : :
QY 572 REKQCADFNDMPFRGKYNNKPYTGGVXP-CALNCLAEVNFYTERAPAVIDGTQCNAD 630
Db : : : : :
QY 745 SRNVICIDIGICERGVCDNLGSDAVEDVCGVGNNSACTIHRGLYTKHHTNQYHVTI 804
Db : : : : :
QY 631 SLDICINGECKHVGCNDILGSDAREDRVCVGGGSGTDAIEGFFNDSLPGRGYMEVVQI 690
Db : : : : :
QY 805 PSGARSIRYEMNVSTYSIVRNALRRYYINGHWTVDMPGRYKFSCTTFDYRRSYNEPEN 864
Db : : : : :
QY 691 PRGSVHLEVRVAMSKNYIALKSEGDYVINGAWTIDWPKPDVAGTAFHYKRPDPEPS 750
Db : : : : :
QY 865 LIATGPTNETLIVELLFQGRNPGVAWEYSMP---RLGTEKQPPAOPSYTWAIVR-SECSVS 921
Db : : : : :
QY 751 LEALGPTSENLIWVLLQSNLGIYKFNVPITRTYSGDN---EVGFTWNHQPMSECSAT 807
Db : : : : :
QY 922 CGGGR 926
Db : : : : :
QY 808 CAGGK 812
Db : : : : :

RESULT 4

US-09-369-364A-7
; Sequence 7, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 997

Query Match 22.5%; Score 1179; DB 4; Length 997;
Best Local Similarity 30.3%; Pred No. 1.5e-96;
Matches 292; Conservative 125; Mismatches 307; Indels 240; Gaps 26;
US-09-369-364A-7

QY 129 DLVSAYEVDHRGDYVSHHEIMHQRRAVAVSEVSLHLRLKGRPHDFHMLRTSSSLVA 188
Db : : : : :
QY 44 DIVHVRVDAGSFLSYELWPRALAKRDVSVRRDAPAFYELQYRGRELRFNLTAQHLLA 103
Db : : : : :
QY 189 PGIIVTCLKTETKSVQTLPPEDFCFYQCSLASHRNSPSHG-----KFCGSTRITLCLCN 244
Db : : : : :
QY 104 PGVSTRRGGGLGRAHITRAHTPACHLLGEV---QDPELEGGLAISACDGLKGVFQLSN 160
Db : : : : :
QY 245 SQKC---PRDSVDFRAAOCAEH-----NSRRFRGRHYKWKPYT---QVEADLCKLYCIAE 293
Db : : : : :
QY 161 EDYFIEPLDSAPARPGHAQPHVVKYKQAPERLAQRGDSSAPSTCGVQVVPCL-----212
Db : : : : :
QY 294 GFDFPFSLSNKVKDGTFCPSDSDSRNVICIDIGICLSVSVSAHMPQPKEDLFIIPDEYKSC 353
Db : : : : :
QY 213 -----ESRR-----ERWSQR 222
Db : : : : :
QY 354 LRHKRSLRS-HR---NEELNVETLVVDKMMQNHGHEHNTYVTLNVMVSAFLKQGL 409
Db : : : : :
QY 223 QWRRRRLRLHORSVSKKWCETLVVADAKVVEYHGQPOQVSVYLTIMNVAGLFHDPS 282
Db : : : : :
QY 410 MGK-----DGT-----RHDHAILLT 424
Db : : : : :
QY 283 IGNPITHTIVRLVLLDEDEEDLKTHTHADNTLKSFKWOKSINMKGDAPHLHDTAILT 342
Db : : : : :
QY 425 GLDICSWKNEPCTDLPAPISGCMKYRSCCTNEDTGLGLAFTTAHESCHNFGMIHDSG 484
Db : : : : :
QY 343 RKDLCAAMRBCETLGLSHVAGMCQPHRSCSNEDTGLPLAFTVAHELCHSFGIQHDSG 402
Db : : : : :
QY 485 NMCKK---SEGNI MSPTLAGNNGVFSWSPCSRQYLHKLFLSTAAQICLADOP-KPVKEYKVP 541
Db : : : : :
QY 403 NDCEPVGKRPFIIMQLLYDAAPLTWSRCSQVITRFLORGWGLCLDDPAPAKDID--FP 460
Db : : : : :
QY 542 EXLPGELYDANTQCKWQFGEKAKLCLMDFKDKICKALMCHRIORCKETKEMPAAGTICG 601
Db : : : : :
QY 461 SVFPGVLYDVSHOCRLQYGAFAFC---EDMDNVCHTLWC-SVGTTCCKSLDAAVDGTRCG 517
Db : : : : :
QY 602 HDWCRGGQCVKYCDEGPKPTHGHSWSSWSPCSRCTCGGVSYSRSLCTNPKPSHGGRF 661
Db : : : : :
QY 518 ENKWLCSGECVPVGPFR-PEAVDGGSGWSAWSICSRSCGNGVQSAERQCTQPTPKYKGY 576
Db : : : : :
QY 662 CEGSTRITLCLNSQKCPROSDVFRAAOCAHNSRFRGRHYKWKPYTQVEDODLCKLYCI 721
Db : : : : :
QY 577 CVGERKXRFRLCNLQACAPGRPSFRHVQCSEHFDAMLYKQLHTWVPV--VNDVNPCELHCR 634
Db : : : : :
QY 722 AEGDFPFSLSNKVKDGTFCPS---DSRNVICIDIGICERGVCDNLGSDAVEDVCGVGNNS 779
Db : : : : :
QY 635 PANEPFAKLRDACVDGTPTCYQVRASTRDLICNGICKNVGCDFEIDSGAMEDRCGVCHNG 694
Db : : : : :
QY 780 SACTIHRGLYTKHHTNQYHVTIPSGARSIRYEMNVSTYSIVRNALRRYYINGHWT 838
Db : : : : :
QY 695 STCHTWSGTPEEAEGLG-YVDVGLIPAGAREINIOEVAEAAFLAIRSDPDKYFLNGW 753
Db : : : : :
QY 839 TVDMPGRYKFSCTTFDYRRSYNEPENLIATGPTNE-----873
Db : : : : :
QY 754 TIQWNGDYVAGTITFTYARGNW-ENLTSFGPTKEPVNIQVPASRGPGGSGRGVPRPST 812
Db : : : : :
QY 874 -----873
Db : : : : :
QY 813 LHGRSRPGGVS PGSVTEPCGSEPGPAAASTSVSPSLKPNLVAHVHGGMGQAPLGLGW 872
Db : : : : :
QY 874 -----TLIVELLFQGRNPGVAWEYSMPRL--GTEKQPPAOPSYTWAIVR-SECSVS 921
Db : : : : :
QY 873 RRLHVLNMGRLPQLTQLLFQSNFNGVHYEYTHIREAGGHDEVPV--PVFWSHWHQFWTKCTVT 930
Db : : : : :
QY 922 CGGK 925

[illegible]

RESULT 12

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

US-09-491-522-5

Query Match
17.6%; Score 923; DB 4; Length 1211;

Query Match	17.6%;	Score 923;	PS 4;	Length 4211,
Best Local Similarity	28.6%;	Pred. No. 1.8e-73;		
Matches	266;	Conservative	132;	Mismatches 363;
		Indels	168;	Gaps 31;

Qy	81	PERPASSSTRGAAGLDGGRMDDEAGNHR	SOQNTWTGTENOTLHVLTQYDLUSVAYEVDHKG	140
Db	29	PPPPPPANARLAAAAPPGGPL-----	-GHGAERI-----	68
Qy	141	DVVSHEIMHQ-----	-RRRAVAV-----	184
Db	69	RUUVSHVGAATSRAGVARRAPVTRTSP	FGCHNEEPGSHLFYNVTVPGEDLHLRLPNA	128
Qy	185	SUVAFFIVQTLGKTGTSKVQTLPPEDPC	FYQVQSLRSHRNSPHSGKFCBGSTRTIKLKN	244
Db	129	RLVAPGATWEMQEGKGTTRVE--	PLLGSLCYVGDVAGLABASSVALSNCGLAGLRMBE	186
Qy	245	SQKCPRDSVDFRAOCCAHNSRRERGR-	HYKM-KPYTQVADLCKLYCAIEGEDFFSLS	302
Db	187	EBFFTEPLEKGLAQAE-----	-QGRVHVYRPPTSPPL-----	221
Qy	303	NKVKOGTFCSEDSRNVICIDGICELSV	SVSTSAHMPQPKEDFILPDEYKSLRHSRLR	362
Db	222	-----GGQALDT-CASLDSJLSL-----	-RALGVLEEHANSSRRRAR--R	259
Qy	363	SHRNELAVETLVVVKKMGNHGHENIT	TVTLNMVSALPKDGLMG-----	411
Db	260	HAADDYNEIVLGVDDSVQFHGKBHQY	KLLTMNIVNEITHDESAGHAINVLVRII	319
Qy	412	-----	-KOGT-----	436
Db	320	LLSYGKMSLBIIGNPSQLENVCRWAY	LOQKPDTHDEYHDAIFLTRQDF-----	374
Qy	437	DTLGAPISGWCVKYRSTINEDTCLG	LAFTIAHESCHNGMTHDGEGMC--	494
Db	375	GMOGAPVTGMCHPYRSTCTLNHEDG	SFAFVAHETHTVLGMEHDGOGNRCGDEVILGSI	434
Qy	495	MSPTIAGRNGVFSMSPCSKQYLRFL	STAQAICLAQPKPVKYKYPEKLPGELYDANTQ	554
Db	435	MAPLVOQAFHRPHWSRCSQQLSR	LYHSYD--CLLDDPF-AHDWPALPOLPGLHYSMNEQ	491
Qy	555	CKWQGEKAKLMLDFKDIKCALWCHRI	GK-CETKFPAAEGTICGHDMWCRGGQCV	612
Db	492	CRFDPLGLYMMCTAFRTTDPCKQL	WCSDHPNPFYCKTKGPPJUGTNCAPGKHCFKGHCI	551
Qy	613	KYDGGKPTGHGHWSDNSMSPCSRT	CGGYSHRSLRCLTNPKPSHGKFCGEGSTRTIKLC	672
Db	552	WLTPDILK-RDGSWGAMSPFGCS	SRTCGTVKFRTRCQDNHPFANGGRTCSGLAYDQLC	610
Qy	673	NSQKCRDUSVDFRAOQA-----	-EHSRRPGRHYANKPYTQVEDDLCKLYCIAEGFD	726
Db	611	SROQDCPSLADFRESCQRMQLYF	EHGDAQ-----	785
Qy	727	FFPFSNKKYKOGTPCS-EDSRNVCID	GICERVGCDNVLGSDAVEDCVGCNGNNSACTIH	785
Db	666	EYVSMKRVVHDTRCSYKDAFSIQ	VRGDCRKVGCDSKQEDKCGVCGGDSHCKVU	725
Qy	786	RLGYTKHHHTNQYHKWTIPSGARS	IIEYKNUVSTSYISVRN-ALRYYLNGHWTVDWEG	844
Db	726	KGTFTRSPKGGYIKMFEIPAGARH	LLOEVDATSHHLAVKNLETKQFILNEENDVDASS	785
Qy	845	R-YKFSGTTDFYRRSYNEPENLIA	TGPTNETLIVELLFQG--RNPGVAMEYSM--	900
Db	786	KTFIANGVENEX-RDEGRETLOTW	GPJHGTITVLVIVGDTRVSTLYKYMHEDSLNDV	844
Qy	901	KOPPAQPS-----	YTWAIUR-SECSVSCGGG	925
Db	845	DNNVLEEDSVVYEWALKNSPC	KPCGGG	873

RESULT 13
US-09-130-491-13

Sequence 13, Application US/09130491
 Patent No. 6416974
 GENERAL INFORMATION:
 APPLICANT: Holzman, Douglas A.
 APPLICANT: Goodearl, Andrew D.J.
 TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
 FILE REFERENCE: 09404/041001
 CURRENT APPLICATION NUMBER: US/09/130,491
 CURRENT FILING DATE: 1998-08-07
 EARLIER APPLICATION NUMBER: US 60/058,108
 EARLIER FILING DATE: 1997-09-05
 EARLIER APPLICATION NUMBER: US 60/054,961
 EARLIER FILING DATE: 1997-08-06
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 13
 LENGTH: 608
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-130-491-13

Query Match 17.2%; Score 903; DB 4; Length 608;
 Best Local Similarity 36.5%; Pred. No. 3.8e-72;
 Matches 206; Conservative 73; Mismatches 236; Indels 50; Gaps 20;

QY 412 KGTGTHDHAILLTGLDICSWKNEPCDTLGFAPISGMSKVSRTINEDTGLGLAFTIAHE 471
 DB 40 RDAEHYDTAILFTREQDLOG--SQTCDTLGWADVCTVCDPSRSCSVIEDDGLQAFITAHE 97

QY 472 SGHNFMTDHEGNNCKKSEG-----NIMPTLAGNGVFSWSPSCRSQYHLKFLSTAQAI 526
 DB 98 LQHVNMPHD--DAKQASLNGVNDSSHMASMLNLDHSPFSCSAYMITSLFDNGHGE 156

QY 527 CLADQPKVKEYPEKLPGLBYDANTCKWOFGEKALCMLDPKDKICKALCHGRIGR 586
 DB 157 CLMD--KQNFQLPGLPGTSYANRQCQFTFEDSKHC--PDASTCSTLNC--TGTS 210

QY 587 -----CETKFWPAAGTTCGHDMCRGQCQVYKGDGEG--PKPTHGHSWDSHSSMSPSCRT 639
 DB 211 GGVLCVQTKHFPAWADGTSCEGSKWCKINGKCNKTKDRKHFDTPFHGSGMGMGPWGDGSR 270

QY 640 GGVSHRSLCTNPKPSHGKFCGSGSTRTKLCNSOKCP--RDSVDFRAAOCAHNN---S 694
 DB 271 GGVGYTWTRECDNPNVPMGKGYCEKGVRYSCMLEOCDDNNGKTFREBOCEAHNEFSKA 330

QY 695 RFRGRHYKWP--YTQVEDQDLCKLYCIAEGDFFFLSLNKKVDGTPCSEDSRNVCTDGI 753
 DB 331 SFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQ 390

QY 754 CERVQCDNVLGSDADEVDCVGCNNGNSACTIHRGLYTKHHHTNOYVHMVTPSGARSIRI 813
 DB 391 CVRAGCDRIIDSKKFKDCKGCGVGGNGSTCKKISGVTS--AKPGYRDIITIPGAINIEV 448

QY 814 YEMNV-----STSYISVARNALRRYLNHGHVTDWPGR--YKPSGTFDYRYSYNPENLIA 867
 DB 449 KQNRQGRNNGSFLAIKAADGTYILNGDYTLSTLEQIMYKGVLYRSGSSAALERIRS 508

QY 868 TQPTMETIIVLLEFQGR--NPGVAWEYSMPRLGTEKQPPAPSVT--WAIYR--SECVSQC 923
 DB 509 FSLKEPLTIQVLTGNALRPKITYFYVKK--KKESEFNAIPFSANVIEENGECSTKCG 566

QY 924 GG-----RCLP--VLLLEAACQP 939
 DB 567 KGVKRSKLKLSHDGGLVSHSCDP 591

RESULT 14

US-09-369-364A-2
 Sequence 2, Application US/09369364A
 Patent No. 6391610
 GENERAL INFORMATION:
 APPLICANT: Apte, Suneel
 APPLICANT: Hurskainen, Tiina L.

APPLICANT: Hirohata, Satoshi
 TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
 FILE REFERENCE: 26473/4007/10-30-00
 CURRENT APPLICATION NUMBER: US/09/369,364A
 CURRENT FILING DATE: 1999-08-06
 NUMBER OF SEQ ID NOS: 31
 SOFTWARE: Patent in Ver. 2.1
 SEQ ID NO 2
 LENGTH: 930
 TYPE: PRT
 ORGANISM: mus musculus ADAMTS-5
 US-09-369-364A-2

Query Match 16.9%; Score 885.5; DB 4; Length 930;
 Best Local Similarity 28.7%; Pred. No. 2.8e-70;
 Matches 273; Conservative 120; Mismatches 354; Indels 205; Gaps 40;

QY 48 LLSRGRRLTAMSPFLS-AGTCVVRHGRSGSAWEPERPASSSTRGAAGLDGKGRDMDEAG 106
 DB 14 LLSASCLSLAADSPAAPAOQKTRQQAASAAAEPPQPOQEETR-----ERGHLOPLAG 67

QY 107 NHRSCQNTGTENOTLHVTQYDLVSAYEVDHHRGDYVSHEIMHHQRRRAVAVSESLH 166
 DB 68 QRRS-----GGLVHNIDQ--LYSG-----GKVGVLVYAGGR----- 98

QY 167 LRLKGRPHDFHMDLRTSSSLVAGFIVOTLGTGTGTSVQTLPPEDFCFYQGSLSRHSRNP 226
 DB 99 -----FLDLERDDTVGAAGSIVTAGGGLSASSGH-----RGHCFYRGTV----- 138

QY 227 SHGKFCFEGSTRTL---KLCNSOKCPRDSVDVFPAAOCAEHNSR--RFRGRHYKWKPYQVE 282
 DB 139 -----DGSPRSLAVFDLQGLD-----GFFAVKHARYTLKPLURG---SWAERYIY 182

QY 283 AD----LCKLYCIAEGDFFFLSLNKKVDGTPCSEDS-----RNVCIDGICE 325
 DB 183 GDGSRILHVT--NREGFS--FEALPPRASCETPASPGQSPQSPSVHSRSLRSLALQILD 240

QY 336 LSVVSTSAHM--POPKKEDFLPDEYKSLRHKRSLRSHRNEELAVETLVVVVDKQMQN 384
 DB 241 HSATSPSGNAGPQ-----TWRRRRRSISRARQ-----VELLVADSSMAR 282

QY 395 HGHENITYVLTILNMVSALP-----KD----- 407
 DB 283 YG--RGLQHYLTMAISIANRLYSHASIHNLAVKVVVLVLTDXDTSLEVSNAATLKNF 341

QY 408 -----GLMGKDOTRH--DHAILLTGDIKSWKNEPCDTLGFAPISGMSKYSCTINED 459
 DB 342 CKWQHQNQLGDDHEEHYDAAILFTREDLGC--HHSCTDLGMADVGTICTSPERSCAVIED 399

QY 460 TGLGLAFTIAHESGHNFGMHDGEGNNCKKSEG-----NIMPTLAGNGVFSWSPSCRSQ 514
 DB 400 DGLHAAPTVAHEIHGHLGLSHD--DSKFCBEENFGTTEDKRLMSSILTSIDASKPWSKCTSA 458

QY 515 YLHFLPLTAQAIADQPKVKEYPEKLPGLBYDANTCKWOFGEKALCMLDPFKKDI 574
 DB 459 TITFLEDDHGNCLELDP--KQILGPEELPGQTYDATQCCLNTFGPEYSVCP---GMDV 513

QY 575 CKALWCH--RIGRK--CETKFWPAAGTTCGHDMCRGQCQVYKGDGEGK-----PTHGHW 626
 DB 514 CARLWCAVVRGQMVCTKLPKLPVAVGTPCCKGRVCLQKCV---DKTKKXYSTSRGNW 570

QY 627 SDNWSWSPSCRTCGGVSHRSLCTNPKPSHGKFCGSGSTRTKLCNSOKCPRDSVDFFRA 686
 DB 571 GSWGPGWQCSRS CGGVQFAVRHNNPAPRNSGRYCTGKEAIYRSCSVTPCPCPNKGSFRH 630

QY 687 AQCAHNSRRRFRGRHYK-----WKP--YTQVEDQDLCKLYCIAEGDFFFLSLNKKVDGTC 741
 DB 631 EQCEAKNGYQSDAKGVKTFVWVPKYAGVLPADVPKLTCTRAKGTGYVYVFPSPKVTDTGTC 690

QY 742 SEDSRNVCTDGI CERVCDNVLGSDADEVDCVGCNNGNSACTIHRGLYTKHHHTNOYVHM 801
 DB 691 RPYNSVVRGRCVTRCTCDGIIIGSKLQYDKCGVCGGDNSSCTKIITPNK---KSKGYTDV 748

[illegible]

Db	S89	FAIRHCNNPARNRNGRYCTGKRALRYRCSLSMPCFPNGKSPFRHCECTACKNGYQSDAKGVKT	548
Qy	704	---WRP-YTQVEDQDLCKLYCIAEGDFFFSLGNKVKDGTGTCSEDSRNVCIDGICERVGC	759
Db	649	FVEMVPKYAGVLPADVCKLATCRAGTGYVYVFPKVTGDTGTCRFPYSNVCVRGKCVRTGC	708
Qy	760	DNLGSDADEVDCGVGNGNSACTIHRGLATKKHHNTQYVHWYT1PSGARSTRIEYMNVS	819
Db	709	DGIIGSKLOYDKCGVCGGDNSCTKIVGTNKK--KSKGYTDVVRIPGATHIKVROFKAK	766
Qy	820	-----TSYISVENALRRYLLNGHWTVDPWGR-KYFSGTTFDYRRSYNEPENIATG--PT	871
Db	767	DQTRFAYLALKKNGEYLNGKVMYISTETIIDIINGVMVNSGWSHRDDFLHGMGSYAT	826
Qy	872	NETLIVELLFOGRNP-----GVANEYSNPRLGTEK-----QPPAPQPSYT--N	911
Db	827	KEILIVQL--ATDPTKPLDVRYSFFVPKKSTPKVNSVTSHGSKNGKVSHTSQPWWTGPW	884
Qy	912	AIVRSECVSVCGGG	925
Db	885	L-----ACSRCTDGT	894

Search completed: October 28, 2003, 23:47:08
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 23:41:00 ; Search time 64 Seconds
(without alignments)
2490.977 Million cell updates/sec

Title: US-09-981-151a-8

Perfect score: 5236
Sequence: 1 MKPRARGRWGLAALWMLLAQ.....LEACQPSATYATALAFLES 952

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCT05_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3846.5	73.5	1224	14	US-10-217-774-4
2	3846.5	73.5	1224	16	US-10-296-616-2
3	2435.5	46.5	1221	12	US-10-240-545A-2
4	2265	43.3	862	15	US-10-226-560-2
5	1543.5	29.5	491	14	US-10-217-774-2
6	1391	26.6	959	10	US-09-788-043C-1
7	1390	26.5	908	10	US-09-963-791-2
8	1376	26.3	684	15	US-10-103-377C-2
9	1319	25.2	757	10	US-09-963-791-24
10	1257	24.0	1690	10	US-09-788-043C-5
11	1243	23.7	1044	9	US-09-858-081-9
12	1240	23.7	1104	12	US-09-842-469-2
13	1234	23.6	1133	9	US-09-858-068-2
14	1223.5	23.3	1068	9	US-09-858-081-2
15	1211	23.1	1104	12	US-09-842-469-4

16	1195	22.8	814	9	US-09-858-081-12	Sequence 12, Appl
17	1188	22.7	859	10	US-09-918-171A-5	Sequence 5, Appl
18	1179	22.5	947	10	US-09-918-171A-7	Sequence 7, Appl
19	1072.5	20.5	997	10	US-09-788-043C-3	Sequence 3, Appl
20	1064.5	20.3	1629	9	US-09-972-467-2	Sequence 2, Appl
21	1063.5	20.3	947	10	US-09-788-043C-7	Sequence 7, Appl
22	1057.5	20.2	1081	10	US-09-918-171A-17	Sequence 17, Appl
23	1056	20.2	1882	10	US-09-918-171A-13	Sequence 13, Appl
24	1045.5	20.0	1907	10	US-09-938-330-25	Sequence 25, Appl
25	1043.5	19.9	1057	12	US-10-188-869-13	Sequence 10, Appl
26	1043.5	19.9	1122	12	US-10-188-869-13	Sequence 13, Appl
27	1029	19.7	1145	12	US-10-188-869-20	Sequence 20, Appl
28	1022.5	19.5	930	10	US-09-321-987B-4	Sequence 4, Appl
29	1022.5	19.5	968	14	US-10-163-316-7	Sequence 7, Appl
30	1018.5	19.5	1213	10	US-09-969-515-8	Sequence 8, Appl
31	1003	19.2	756	12	US-10-188-869-6	Sequence 6, Appl
32	1003	19.2	1224	10	US-09-969-515-2	Sequence 1, Appl
33	998	19.1	1205	15	US-10-164-890-1	Sequence 1, Appl
34	991	18.9	874	10	US-09-918-171A-15	Sequence 15, Appl
35	988.5	18.9	779	12	US-10-188-869-8	Sequence 8, Appl
36	984.5	18.8	950	11	US-09-373-658-2	Sequence 2, Appl
37	984.5	18.8	967	12	US-10-115-286-2	Sequence 2, Appl
38	984.5	18.8	967	14	US-10-105-929-2	Sequence 2, Appl
39	984.5	18.8	968	11	US-09-373-658-125	Sequence 125, Appl
40	979.5	18.7	969	10	US-09-969-515-10	Sequence 10, Appl
41	969	18.5	2150	10	US-09-321-987B-2	Sequence 2, Appl
42	965.5	18.4	2165	9	US-09-800-729-155	Sequence 155, Appl
43	964	18.4	980	10	US-09-969-515-4	Sequence 4, Appl
44	958.5	18.3	727	10	US-09-445-023A-12	Sequence 12, Appl
45	958.5	18.3	727	15	US-10-097-597-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-10-217-774-4
; Sequence 4, Application US/10217774
; Publication No. US20020193583A1
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020193583A1 Human Proteases and Polynucleotides Encod
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-217-774-4

Query Match	73.5%	Score	3846.5	DB	14	Length	1224
Best Local Similarity	75.4%	Pred. No.	0				
Matches	738	Conservative	25	Mismatches	73	Indels	143
Gaps	13						
QY	1	MKPRARGRWGLAALWMLLAQVAEVSFGSRSHQRNRCGQLEASPPRLSRGPRRLTAMS	60				
DB	1	MKPRARGRWGLAALWMLLAQVAEAP	31				
QY	61	PLFESAGTCVRHGTSGGAWEPERASSSTRGACLDGKGRDMDAGNHRSCQNTGTENQ	120				
DB	32	PAAAA	58				
QY	121	TLHLVTQYDLVSAYEVDHGRGYSHEIMHHQRRRAVAVSEVSLHLRLKGRPHDPMFL	180				


```

Db 59 -----EYDLVSAVEVDHGRDYVSHEIMHQRRAVAVSEVESLHLRLKGRSHDFHVDL 112
Qy 181 RTSSSLVAPGFIQVTLGKTGTSVOTLPDEDFCFYQGLSRSHNSPSHGKFCGSTRTL 240
Db 113 RTSSSLVAPGFIQVTLGKTGTSVOTLPDEDFCFYQGLSRSHNS-SVALSTCQGLSGMI 171
Qy 241 KLCNSQKPRDSVDFRAAQCAEHSNRRFRGR-----HYKWKPYTOVEA-DLCKLYCI 291
Db 172 R-----TEADYFLRPLPSHLGWL-GRAAQSSPSHVLKRSSTEPHAPGASEVLVT 222
Qy 292 AEGFDFEFS--LSNKKVDGTPCEDSRNVCIDGICELSVVTSVAHMPQPPKEDLILPDE 349
Db 223 SRTWELAHQPLHSSDLRLGLP---QKQHFC-----GRRKKYMPQPPKEDLILPDE 270
Qy 350 YKSLRHKRSLLSRHNEELNVTLLVVDKQKQNHGHNITTYVLTILNMVSALFKD-- 407
Db 271 YKSLRHKRSLLSRHNEELNVTLLVVDKQKQNHGHNITTYVLTILNMVSALFKDGT 330
Qy 408 -----GLMGKDGTRHDHAILLTGLD 427
Db 331 IGGNINIAVGLILLEDEQPLVISHHADHTLSSFCQWQSLMGKDGTRHDHAILLTGLD 390
Qy 428 ICSWKNEPCDTLGFAPISGMSKYSRCTINEDTGLGFLAFTIAHESGHNFMIHDGEGNMC 487
Db 391 ICSWKNEPCDTLGFAPISGMSKYSRCTINEDTGLGFLAFTIAHESGHNFMIHDGEGNMC 450
Qy 488 KKSSEGNIMSPTLAGRNQVFSWSPCSROYLHKFLSTAAICLADQPKVKEYKYPEKLPG 547
Db 451 KKSSEGNIMSPTLAGRNQVFSWSPCSROYLHKFLSTAAICLADQPKVKEYKYPEKLPG 510
Qy 548 LYDANTOCKWQFGEKAKLMDPKDICKALWCHRIGRKCEKTFMPAAEGTICGHDMWCR 607
Db 511 LYDANTOCKWQFGEKAKLMDPKDICKALWCHRIGRKCEKTFMPAAEGTICGHDMWCR 570
Qy 608 GGQCVKYGDEGPKPTHGHSNDSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 667
Db 571 GGQCVKYGDEGPKPTHGHSNDSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 630
Qy 668 TLKLCNSQKPRDSVDFRAAQCAEHSNRRFRGRHYKWKPYTOVEDDLCKLYCIAEGDF 727
Db 631 TLKLCNSQKPRDSVDFRAAQCAEHSNRRFRGRHYKWKPYTOVEDDLCKLYCIAEGDF 690
Qy 728 FFSLSNKVKDGTPCSEDSRNVCIDGICERVCNDNLGSDADEVCGVCNNSACTIHRG 787
Db 691 FFSLSNKVKDGTPCSEDSRNVCIDGICERVCNDNLGSDADEVCGVCNNSACTIHRG 750
Qy 788 LYTKHHHTNQYHHMTIPSGARSIRIYEMNVSTYSVRNALRRYLLNGHWTVDWPGRYK 847
Db 751 LYTKHHHTNQYHHMTIPSGARSIRIYEMNVSTYSVRNALRRYLLNGHWTVDWPGRYK 810
Qy 848 FSGTTFDYRRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQ 907
Db 811 FSGTTFDYRRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQ 870
Qy 908 SYTMAIVRSECSVSCGGGR 926
Db 871 SYTMAIVRSECSVSCGGG 889

```

RESULT 2

```

US-10-296-616-2
; Sequence 2, Application US/10296616
; Publication No. US20030129658A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: No. US20030129658A1el protease
; FILE REFERENCE: Y0132PCT-664
; CURRENT APPLICATION NUMBER: US/10/296.616
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: JP 2000-393372
; PRIOR FILING DATE: 2000-12-25
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 2

```

```

; LENGTH: 1224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-616-2

Query Match 73.5%; Score 3846.5; DB 16; Length 1224;
Best Local Similarity 75.4%; Pred. No. 0;
Matches 738; Conservative 25; Mismatches 73; Indels 143; Gaps 13;

Qy 1 MKPRARGWRCGLAALWMLLAQVAVSQSPGRSHQRNGRSGQLBASPPRLLSRGRPRLTAMS 60
Db 1 MKPRARGWRCGLAALWMLLAQVAVSQSPGRSHQRNGRSGQLBASPPRLLSRGRPRLTAMS 60
Qy 61 PLFSAGTCVRHGRTRSGSAMPEPASPSTRTGAAGLDGKGRDMDAGNHSQQNTNTENQ 120
Db 32 PAAAA-----PGSPSVPRPPPAERPG-----MMKRG-----58
Qy 121 TLHVLTOYDLVSAVEVDHGRDYVSHEIMHQRRAVAVSEVESLHLRLKGRSHDFHMDL 180
Db 59 -----EYDLVSAVEVDHGRDYVSHEIMHQRRAVAVSEVESLHLRLKGRSHDFHVDL 112
Qy 181 RTSSSLVAPGFIQVTLGKTGTSVOTLPDEDFCFYQGLSRSHNSPSHGKFCGSTRTL 240
Db 113 RTSSSLVAPGFIQVTLGKTGTSVOTLPDEDFCFYQGLSRSHNS-SVALSTCQGLSGMI 171
Qy 241 KLCNSQKPRDSVDFRAAQCAEHSNRRFRGR-----HYKWKPYTOVEA-DLCKLYCI 291
Db 172 R-----TEADYFLRPLPSHLGWL-GRAAQSSPSHVLKRSSTEPHAPGASEVLVT 222
Qy 292 AEGFDFEFS--LSNKKVDGTPCEDSRNVCIDGICELSVVTSVAHMPQPPKEDLILPDE 349
Db 223 SRTWELAHQPLHSSDLRLGLP---QKQHFC-----GRRKKYMPQPPKEDLILPDE 270
Qy 350 YKSLRHKRSLLSRHNEELNVTLLVVDKQKQNHGHNITTYVLTILNMVSALFKD-- 407
Db 271 YKSLRHKRSLLSRHNEELNVTLLVVDKQKQNHGHNITTYVLTILNMVSALFKDGT 330
Qy 408 -----GLMGKDGTRHDHAILLTGLD 427
Db 331 IGGNINIAVGLILLEDEQPLVISHHADHTLSSFCQWQSLMGKDGTRHDHAILLTGLD 390
Qy 428 ICSWKNEPCDTLGFAPISGMSKYSRCTINEDTGLGFLAFTIAHESGHNFMIHDGEGNMC 487
Db 391 ICSWKNEPCDTLGFAPISGMSKYSRCTINEDTGLGFLAFTIAHESGHNFMIHDGEGNMC 450
Qy 488 KKSSEGNIMSPTLAGRNQVFSWSPCSROYLHKFLSTAAICLADQPKVKEYKYPEKLPG 547
Db 451 KKSSEGNIMSPTLAGRNQVFSWSPCSROYLHKFLSTAAICLADQPKVKEYKYPEKLPG 510
Qy 548 LYDANTOCKWQFGEKAKLMDPKDICKALWCHRIGRKCEKTFMPAAEGTICGHDMWCR 607
Db 511 LYDANTOCKWQFGEKAKLMDPKDICKALWCHRIGRKCEKTFMPAAEGTICGHDMWCR 570
Qy 608 GGQCVKYGDEGPKPTHGHSNDSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 667
Db 571 GGQCVKYGDEGPKPTHGHSNDSWSPCSRTCCGGVSHRSRLCTNPKPSHGKFCGSTR 630
Qy 668 TLKLCNSQKPRDSVDFRAAQCAEHSNRRFRGRHYKWKPYTOVEDDLCKLYCIAEGDF 727
Db 631 TLKLCNSQKPRDSVDFRAAQCAEHSNRRFRGRHYKWKPYTOVEDDLCKLYCIAEGDF 690
Qy 728 FFSLSNKVKDGTPCSEDSRNVCIDGICERVCNDNLGSDADEVCGVCNNSACTIHRG 787
Db 691 FFSLSNKVKDGTPCSEDSRNVCIDGICERVCNDNLGSDADEVCGVCNNSACTIHRG 750
Qy 788 LYTKHHHTNQYHHMTIPSGARSIRIYEMNVSTYSVRNALRRYLLNGHWTVDWPGRYK 847
Db 751 LYTKHHHTNQYHHMTIPSGARSIRIYEMNVSTYSVRNALRRYLLNGHWTVDWPGRYK 810
Qy 848 FSGTTFDYRRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQ 907
Db 811 FSGTTFDYRRSYNEPENLIATGPTNETLIIVELLFQGRNPGVAWEYSMPRLGTEKQPPAQ 870

```


QY 908 SYTWAIVRSECVSCGGG 926
Db 871 SYTWAIVRSECVSCGGG 889

RESULT 3

US-10-240-545A-2
; Sequence 2, Application US/10240545A
; Publication No. US20030185828A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: No. US20030185828A1el aggrcanase
; FILE REFERENCE: 08959.0002
; CURRENT APPLICATION NUMBER: US/10/240.545A
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/JP01/11033
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: JP 2000-384300
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 2
; LENGTH: 1221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-545A-2

Query Match 46.5%; Score 2435.5; DB 12; Length 1221;
Best Local Similarity 51.3%; Pred. No. 1.3e-207;
Matches 466; Conservative 118; Mismatches 195; Indels 129; Gaps 17;

QY 85 ASSSTRGAAGLDGKGRDMEAGNHRSQQTNTGTENQTLHLVLTQYDLVSAYEVDHRRGDYVS 144
Db 47 ASDSSGASGLN-----DDYVFTVPEVDSAGSYIS 77
QY 145 HEIHHQRRRAVAVSEVESLHLKGRPHDFHMDLTSSSLVAPGFIQTLKTKTSV 204
Db 78 HDILHNGKRS-AQNARSSLHYRFSAGQELHLEK-PSAILSSHFIVQVLGNDGASET 135
QY 205 QTLPPEDFCFYQGSRLSRHNSPSHGKFCGSGTRTLKLCNSQ-----KCPDRSDVDFRAAQC 260
Db 136 QK-PEVQCFCYQGIIRND-SSSSVAVSTCAGLSGLIRKNEFLISPLPQ-----LLA 186
QY 261 AEHNSRRRGRHYKWKPYTQVEADLCKLYCIAEGDFDFSLSNKVKDG-----TPCSEDS 315
Db 187 QEHNYSPAG-HHPHVLKRYTAEKIQRYGPG-----SGRNVPGYSPSHIPASQS 238
QY 316 R-----NVCIDGICELSVSTSAHMPQPPKEDFLIPDEYKSLRHKRSLRS 363
Db 239 RETEYHRRLOKQHF-----GRRKKYAPKPTEDTYLRFDEYSGSRPRRSAGKS 289
QY 364 HRNEELVETLVVVDKXKQKQKHENITTYVLTILNMVSLFKDG----- 408
Db 290 QKG--LNVTLLVADKXKVEKHGKGNVTYILTVMNVSGLFXDGTIGSDINVVVSLIL 347
QY 409 -----LMGKDGTRHDHAILLTGLDTCMKNEPCDTLGF 441
Db 348 LEQEGGLINHHADQSLNSFCQMSALIGKNGKRHDHAILLTGFDICSKWNEPCDTLGF 407
QY 442 APIGMCCKYKSCNTNEDTGLAFTTAHSGHNFHGHGECNMCKSGENIMSPFLAG 501
Db 408 APIGMCCKYKSCNTNEDTGLAFTTAHSGHNFHGHGECNMCKSGENIMSPFLAG 467
QY 502 RNVFMSFCRSQVHLKFLSTAQAIICLADQPKPKVYKYPEKLPGLDYDANTOCKWQFGE 561
Db 468 NNGVFSNCSRSQYLKXELSTFQACCLVDEPKQAGYKPKLPGQIYDADTOCKWQFGE 527
QY 562 KAKLCMLDFKXDKALMCHIRKCTKFNPAEAGTICGHDMWCRGQCVKYVDEGPKP 621
Db 528 KAKLCMLDFKXDKALMCHIRKCTKFNPAEAGTICGHDMWCRGQCVKYVDEGPKP 587
QY 622 THGMSDWSSNPSCSRTCGGVSHRSLCTNPKPSHGKFCGSGTRTLKLCNSQKCPROS 681
Db 588 IHGQSAWSKSECSRTCGGVSHRSLCTNPKPSHGKFCGSGTRTLKLCNSQKCPROS 647

QY 682 VDFRAAQCAEHNRRRGRHYKWKPYTQVEADLCKLYCIAEGDFDFSLSNKVKGTGTPC 741
Db 648 LDFRAQQCAEYNSKPFRCGWFYQWPKYTKVEEDRCKLYCKAENPEFFPFAMSGKVKGTGTPC 707
QY 742 SEDSRNVICDIGICERVGCDNLGSDAIVEDVCGVCGNGNSACTIHRGLYTKHHHTNOYYHM 801
Db 708 SPKNVDVICDGVCELVGCDHGLSKAVSDACGVCKGDNSTCKFYKGLYLNOHRANEYFV 767
QY 802 VTIPSGARSIRIYEMNVSTSYISVRNALRRYLLNGHWTVDWPGRYKFSGTTFDYRRSYNE 861
Db 768 VIIIPAGARSIEIQELQVSSSYLAUVRSLSQKYLGTGWSIDWPGRFPFAGTTFEYQSRFNR 827
QY 862 PENLIATGPTNETLIVELLFQGRNPGVAVWEYSMPRL--GTEKOPPA--QPSYTWAIVRSE 917
Db 828 PERLYAPGPTNETLIVFBIHQGNKPGIAWKYALPKVNMGT---PPATKRFPAYTMSIVQSE 884
QY 918 CSVSCGGG 925
Db 885 CSVSCGGG 892

RESULT 4

US-10-226-560-2
; Sequence 2, Application US/10226560
; Publication No. US20030050464A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Waiker, D. Wade
; TITLE OF INVENTION: No. US20030050464A1el Human Proteases and Polynucleotides Encod

; FILE REFERENCE: LEX-0372-USA
; CURRENT APPLICATION NUMBER: US/10/226.560
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US 60/314,049
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 09/917,614
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/221,644
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 862
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-226-560-2

Query Match 43.3%; Score 2265; DB 15; Length 862;
Best Local Similarity 50.7%; Pred. No. 1.2e-192;
Matches 435; Conservative 107; Mismatches 194; Indels 122; Gaps 14;

QY 85 ASSSTRGAAGLDGKGRDMEAGNHRSQQTNTGTENQTLHLVLTQYDLVSAYEVDHRRGDYVS 144
Db 47 ASDSSGASGLN-----DDYVFTVPEVDSAGSYIS 77
QY 145 HEIHHQRRRAVAVSEVESLHLKGRPHDFHMDLTSSSLVAPGFIQTLKTKTSV 204
Db 78 HDILHNGKRS-AQNARSSLHYRFSAGQELHLEK-PSAILSSHFIVQVLGNDGASET 135
QY 205 QTLPPEDFCFYQGSRLSRHNSPSHGKFCGSGTRTLKLCNSQ-----KCPDRSDVDFRAAQC 260
Db 136 QK-PEVQCFCYQGIIRND-SSSSVAVSTCAGLSGLIRKNEFLISPLPQ-----LLA 186
QY 261 AEHNSRRRGRHYKWKPYTQVEADLCKLYCIAEGDFDFSLSNKVKDG-----TPCSEDS 315
Db 187 QEHNYSPAG-HHPHVLKRYTAEKIQRYGPG-----SGRNVPGYSPSHIPASQS 238
QY 316 R-----NVCIDGICELSVSTSAHMPQPPKEDFLIPDEYKSLRHKRSLRS 363
Db 239 RETEYHRRLOKQHF-----GRRKKYAPKPTEDTYLRFDEYSGSRPRRSAGKS 289

Db 499 CALNCLAEYNYFYTERAPAVIDGTQCNADSLDINGECKHVGCDNIGLSDAREDRVCV 558
 QY 776 NGNSACTIHRGLYTKHHHTNQYHYMYTIPSGARSIRIVEMVSTSYISVBNALRRYLN 835
 Db 559 GGDSTDAIEGFNDLSLPRGYNVEVQIIRGSHVIEVREVMKRYIALKSEGDDYIN 618
 QY 836 GHWTVDMGRYKFGSTFDVRRSNEPENLIATGPNETLIVELLQGRNPGVAMYSMP 895
 Db 619 GAWTIDWPKEDVAGTAHYKRYPTDEPESLEALGPTSENLIWMLVLLQEQNLGIRYFNPV 678
 QY 896 --RLGTEKQPPAQPSTYMAIVR--SECSVSCGGGR 926
 Db 679 ITRIGSDN---EVGFNWHOPWSECSATCAGK 709

 RESULT 10
 US-09-788-043C-5
 ; Sequence 5, Application US/09788043C
 ; Patent No. US20020107361A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heller, Renu
 ; APPLICANT: Zuo, Fengrong
 ; APPLICANT: Klonowski, Paul
 ; TITLE OF INVENTION: No. US20020107361A1el Metalloproteases Having
 ; TITLE OF INVENTION: Thrombospondin Domains and Nucleic Acid Compositions
 ; TITLE OF INVENTION: Encoding the Same
 ; FILE REFERENCE: ROCH-004
 ; CURRENT APPLICATION NUMBER: US/09/788,043C
 ; CURRENT FILING DATE: 2001-02-16
 ; PRIOR APPLICATION NUMBER: 60/184,152
 ; PRIOR FILING DATE: 2000-02-18
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 1690
 ; TYPE: PRT
 ; ORGANISM: human
 US-09-788-043C-5

Query Match 24.0%; Score 1257; DB 10; Length 1690;
 Best Local Similarity 34.1%; Pred. No. 2.66-102;
 Matches 293; Conservative 128; Mismatches 313; Indels 126; Gaps 25;

 QY 129 DLVSAYEVDRGDIYSHIMHQRRAVAVESVLSHLKGRPHDFHDLTSSSLVA 188
 Db 48 DIIVRVYDAGGSFSLSYELMPLRALRKDVSVRRDAPAFYBLQYRGRELRLNLTANQHLLA 107
 QY 189 PGFIVTGLKGTGKSVQTLPPDFCFYQGSLSRHSRSPSHGG---KFCGEGSTRTLKLCN 244
 Db 108 RGFVSETRRRGGLGRAHRAHTPACHLLGEV---QDPELEGGLAAISACDGLKGVQLSN 164
 QY 245 SOKC--PRDSVDFRAAQCAEHSRFRGRHYKRYPTQVEADICKLYCIAEGDFPFSL 302
 Db 165 EDYFTIEPLDSAPAPGHAQPHV-----YKQOAPER-----LA 197
 QY 303 NKVKDGTGTPCEDSRNVCIDGICELSVSTSAHMPQPKEDFILPDEYKSLRHKRSLLR 362
 Db 198 QRGDSAP-----STCGVQVYPELEPRERN-----RQOQWRPR-LRR 236
 QY 363 SHR---NBLNVETLVVVKKQMKQNHGENTITTVTLNWSALFKDGLMGK-----412
 Db 237 LHQRSVSKWKVETLVADAKMVEYHGQPVESYVLTIMNVAAGLHDPDSIGNPIHITIV 296
 QY 413 -----DGT-----RHDHAILLTGLDICSWE 434
 Db 297 RLVLLEDREEDLKIITHADNTPKSCWKQKSIKMKGDANPLHDTAILLTLEKDLCAHMR 356
 QY 435 PCDTGLFAPISGMSKYSRSTINEDTGLGLAFTTAHSGNHPGMIHDEGNMCK--SEG 492
 Db 357 PCDTGLSHVAGMCPHRSCTINEDTGLPLAFTVAHELHSGFGIQHDSGNDCEPVGKRP 416
 QY 493 NIMSPYLAGENGVSFSPCSRYLHKFLSTQAICLADQP-KPVKEYKPEKLPGLYDA 551

Db 417 FIMSPQLLYDAAPLTWSRCSRQYITRFLDRGWSGLCLDDPPAKDID--PPSVPPGVLYDV 474
 QY 552 NTQCKWQFGEKAKLCLMLDFKKOICKALWCHRIKRCETKFMPAABGTICGHDMWCRGGQC 611
 Db 475 SHQCRLOYGAYS AFC--EDMDNVCHTLAC-SVGTTCCHSKLDAADVSTRCENKACLSGEC 531
 QY 612 VKYDDEPKETHGHWDSSWSCSPSTCGGVSHSRRLCTNPKPSHGGKFCGSGTTLKL 671
 Db 532 VPVGRF--PERAVDGGWSGWSAWSTCSRSCGNGVQSAERQCTQTPKYKRYCYGGERKRRFL 590
 QY 672 CNSOKCPEDSVDFRAAQCAEHSRFRGRHYKWKPKYQVEDODLCKLYCIAEGDFPFSL 731
 Db 591 CNLQACAPAGPSPFRHVQCSHFDAWLYKGRLHTWVPV--VNDVNECELHCRPANEYFAEK 648
 QY 732 SNKVKDGTGTPCE--DSRNVCIDGICERVCNDVNLGSDAIVEDVCGVCNGNNSACTIHRGLY 789
 Db 649 RDAVVDGTPCYQVRAGRDLCLNGICRNVGCDPFDISGAMEDRCGVCNGSGTCHTVSGTF 708
 QY 790 TKHHHTMYHYMYTIPSGARSIRIYEMNVTSTSYISVNA-LRRYYLNGHWTVDWPGRYKF 848
 Db 709 EEAELG-YVDVGLIPAGAREIRIQEVAEAAFLALRSEDPEKYFLNGGWTQIMNGDYQV 767
 QY 849 SGTTFDYRSYNEPENLIATGPNETLIVELLQGRNPGVAMYSMPRL--GTEKQPPAQ 906
 Db 768 AGTTFTYARRGNW-ENLTSPTKEPVYIQLLFQESNPGVHYEYTIHREAGGHDEVPP--824
 QY 907 PSYTMAL-VRSECSVSCGGG 925
 Db 825 PVFSWHYGPWTKTIVTCGRG 844

 RESULT 11
 US-09-858-081-9
 ; Sequence 9, Application US/09858081
 ; Patent No. US20020072490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kapeller-libermann, Rosana
 ; APPLICANT: Cook, William James
 ; APPLICANT: Silos-Santiago, Immaculada
 ; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
 ; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
 ; FILE REFERENCE: 10448-049001
 ; CURRENT APPLICATION NUMBER: US/09/858,081
 ; CURRENT FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: 60/204,159
 ; PRIOR FILING DATE: 2000-05-15
 ; PRIOR APPLICATION NUMBER: 60/204,160
 ; PRIOR FILING DATE: 2000-05-15
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 1044
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-858-081-9

Query Match 23.7%; Score 1243; DB 9; Length 1044;
 Best Local Similarity 31.4%; Pred. No. 2.2e-101;
 Matches 295; Conservative 133; Mismatches 317; Indels 194; Gaps 24;

 QY 84 PASSTRGAAGLQGRMDDEAGNHRSSQQTNTGTENQTLHLVTQYDLVSAYEVDRHGDYV 143
 Db 3 PACQILWALAL-GLGLMFEVTHAFRSQ-----DEFLSSLESYEIAFPTRVDHNGALL 54
 QY 144 SHEIMHQRRAVAVESVLSHLKGRPHDFHMDLRSSSIVAPCFIVQILGKTGTS 203
 Db 55 AFSPPPPRQRGRGTATBSRLFYKVASPTSHLLNLTSSRLLAGHVSVEVWTRGLAM 114
 QY 204 VQTLPEPDFCFYQGSLSRHSRNS-----PSHGKFCGSGTR 238
 Db 115 QRAARPH--CLYAGHLQQAQASSSHVAISTCGLHGLIADVEEYLYLEPLHGGP--KGS--168
 QY 239 TLKCNQKCPDRS---VDFRAAQ-----CAEHSRFRGRHYKWKPKYQVEADLC 286


```

Db 169 -----RSPESGPHVYKSSLRHPLDITACGVDRDEKPKWGRPWMLRTLKPPPAR-- 210
Qy 287 KLYCIAEGDFPFFSLNKVKDGTGTPCEDSRNVCIDICELSVSVTSAAHPQPKEDLFL 346
Db 219 -----PLGNETERGQP-----GL----- 231
Qy 347 PDEYKSLRHRKRLSLRSHRNEELNVLTVVVDKMMQNHGHENITTYVLTILNMVSAFLK 406
Db 232 -----KRSVSR-----ERYVETLVVADKMWVAYHGRDVEQYVLAINNIVAKLFQ 276
Qy 407 DGLMG-----KDG- 414
Db 277 DSSLSGTVNLTVRLILLTDOPTLEITHAGKSLDSFCKWQKSI VNHSGHGNAPENG 336
Qy 415 TRHDHAILTGLDICSWKNEPCDTLGFAPISGCMKSVKSCYRSCINEDTGLGLAFTIAHESGH 474
Db 337 ANHDTAVLITRYDICIYKVKPCGTGLGAPVGCRCERSCSVNEDIGLATAFTIAHESGH 396
Qy 475 NFGMIHDEGQWMC---KSEGNIMSPTLAGNMGVFSWSPCSROYLHKFLSTQAICLADQ 531
Db 397 TFGMNHGCVGNSCGARGQDPKALMAAHTMKTNPFFVWSSCSRDYITPSLDSGLGLCLNRR 456
Qy 532 PKPVKEYKPEKLPGLYDANTOCKWQFGEKAKLCLMDFKDOI CKALWCHRIKCKETKF 591
Db 457 P-PRQDFVYTVAPGQAYDADEQCRFQHGKVSQCKYK---EVCSELWCLSKSNRCITNS 512
Qy 592 MPAAGETIC-GHDM---WCRGGQCVKYGDEGPKPTHGHMSDSSWSPCSRTCGGVSHRS 647
Db 513 IPAAEGTLCTHTIDKMCVKYKVPFGSR--PSGVDGAWGMPWTPWDCSRTCGGVSSSS 571
Qy 648 RLCTNPKPSGGKFCESGTRTLKLCNSOKCPDSDVDFRAAQCAEHNSRRFRGRHYKWKPY 707
Db 572 RHCDSPRTTGGKYCLGERRRHRSNTDDCPGSGQDFREVQCSFDSIIFRGRFYKWKTY 631
Qy 708 TVQEDQDLKLYCIAEGDFPFFSLNKVKDGTGTPCEDSRNVCIDICERVCNDVGLGSDA 767
Db 632 -RGGGVKACSLTCLAEGFNFYTERAAVVDGTPCRPDVTDICVSGECKHWGCDRLVGLSD 690
Qy 768 VEDVCGVCGNNSACTIHRGLYTKHHHTNOYHVMVTIPSGARSIRIYEMNVSTSVISVRN 827
Db 691 REDKRCVCGGDSACETIEGVFSPASPGAGYEDVWIPKGSVHIQDNLNLSHLALKG 750
Qy 828 ALRRYVNLGHWTVDWPGYKFSGTTFDYRRSYNEPENLIATGPTNLTVELLFOGRNPG 887
Db 751 DQESLLEGLPGTGPQPHRLPLAGTTTQLRQGPQVQSLEALGPINASLIVMLVARTELPA 810
Qy 888 VAMEYSNPRLGTEKQPPAQPSYTWAIVR-SECSVSCGGG 925
Db 811 LRYRFNAP-IARDSLPP-----YSHWYAPWTKCSAQACAGG 844

```

RESULT 12

```

US-09-842-469-2
; Sequence 2, Application US/09842469
; Publication No. US20030166899A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER PRODUCT INC
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PC10873B
; CURRENT APPLICATION NUMBER: US/09/842,469
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1104
; TYPE: PRT
; ORGANISM: amino acid/human
US-09-842-469-2

```

```

Query Match 23.7%; Score 1240; DB 12; Length 1104;
Best Local Similarity 31.3%; Pred. No. 4.5e-101;

```

```

Matches 294; Conservative 134; Mismatches 317; Indels 194; Gaps 24;
Qy 84 PASSSTGAAGLQCKGRDMDACNHHSSQQTNTGTENQTLVLTQYDLVSAVEYDHRGDYV 143
Db 3 PACOILKWAAL-GLGLMFEVTHAFRSQ-----DEFLSSLESYEIAETPRVDHNGAL 54
Qy 144 SHEIMHQRARRAVAVSEVSLRLRLKGRPHRDMRLTSSSLVAFGFTVQTLGKTGTS 203
Db 55 AFSPPPRRQRGTGATAESRLFYKVASPSTHPLNLTSSRLLAGHVSVEYVWTRGLAM 114
Qy 204 VQTLPPEDFCFYQGSLSHRNS-----PSHGKFCGSTR 238
Db 115 QRAARPH--CLYAGHLQQAQASTSHVAISTCGGLHGLIADBEYELIEPLHGGP--KGS-- 168
Qy 239 TLKLCNSQKCPDSD--VDFRAAQ-----CAHNSRRFRGRHYKWKPYQVQEAELC 286
Db 169 -----RSPESGPHVYKSSLRHPLDITACGVDRDEKPKWGRPWMLRTLKPPPAR-- 218
Qy 287 KLYCIAEGDFPFFSLNKVKDGTGTPCEDSRNVCIDICELSVSVTSAAHPQPKEDLFL 346
Db 219 -----PLGNETERGQP-----GL----- 231
Qy 347 PDEYKSLRHRKRLSLRSHRNEELNVLTVVVDKMMQNHGHENITTYVLTILNMVSAFLK 406
Db 232 -----KRSVSR-----ERYVETLVVADKMWVAYHGRDVEQYVLAINNIVAKLFQ 276
Qy 407 DGLMG-----KDG- 414
Db 277 DSSLSGTVNLTVRLILLTDOPTLEITHAGKSLDSFCKWQKSI VNHSGHGNAPENG 336
Qy 415 TRHDHAILTGLDICSWKNEPCDTLGFAPISGCMKSVKSCYRSCINEDTGLGLAFTIAHESGH 474
Db 337 ANHDTAVLITRYDICIYKVKPCGTGLGAPVGCRCERSCSVNEDIGLATAFTIAHESGH 396
Qy 475 NFGMIHDEGQWMC---KSEGNIMSPTLAGNMGVFSWSPCSROYLHKFLSTQAICLADQ 531
Db 397 TFGMNHGCVGNSCGARGQDPKALMAAHTMKTNPFFVWSSCSRDYITPSLDSGLGLCLNRR 456
Qy 532 PKPVKEYKPEKLPGLYDANTOCKWQFGEKAKLCLMDFKDOI CKALWCHRIKCKETKF 591
Db 457 P-PRQDFVYTVAPGQAYDADEQCRFQHGKVSQCKYK---EVCSELWCLSKSNRCITNS 512
Qy 592 MPAAGETIC-GHDM---WCRGGQCVKYGDEGPKPTHGHMSDSSWSPCSRTCGGVSHRS 647
Db 513 IPAAEGTLCTHTIDKMCVKYKVPFGSR--PSGVDGAWGMPWTPWDCSRTCGGVSSSS 571
Qy 648 RLCTNPKPSGGKFCESGTRTLKLCNSOKCPDSDVDFRAAQCAEHNSRRFRGRHYKWKPY 707
Db 572 RHCDSPRTTGGKYCLGERRRHRSNTDDCPGSGQDFREVQCSFDSIIFRGRFYKWKTY 631
Qy 708 TVQEDQDLKLYCIAEGDFPFFSLNKVKDGTGTPCEDSRNVCIDICERVCNDVGLGSDA 767
Db 632 -RGGGVKACSLTCLAEGFNFYTERAAVVDGTPCRPDVTDICVSGECKHWGCDRLVGLSD 690
Qy 768 VEDVCGVCGNNSACTIHRGLYTKHHHTNOYHVMVTIPSGARSIRIYEMNVSTSVISVRN 827
Db 691 REDKRCVCGGDSACETIEGVFSPASPGAGYEDVWIPKGSVHIQDNLNLSHLALKG 750
Qy 828 ALRRYVNLGHWTVDWPGYKFSGTTFDYRRSYNEPENLIATGPTNLTVELLFOGRNPG 887
Db 751 DQESLLEGLPGTGPQPHRLPLAGTTTQLRQGPQVQSLEALGPINASLIVMLVARTELPA 810
Qy 888 VAMEYSNPRLGTEKQPPAQPSYTWAIVR-SECSVSCGGG 925
Db 811 LRYRFNAP-IARDSLPP-----YSHWYAPWTKCSAQACAGG 844

```

RESULT 13

```

US-09-858-068-2
; Sequence 2, Application US/09858068
; Patent No. US2002007678A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Liebermann, Rosana

```


